

HIV Database Workshop

www.hiv.lanl.gov

seq-info@lanl.gov

Presenters:

Will Fischer, Brian Foley, Bette Korber

Database Pls: Bette Korber, Thomas Leitner,
Karina Yusim, Brian Foley

Additional database staff: Werner Abfalterer,
Kumkum Ganguly, Jennifer Macke, James Szinger, and Hyejin Yoon

Contract Officer Representative: Anjali Singh, NIAID, NIH

*Theoretical Biology and Biophysics, T-6
Los Alamos National Laboratory*



Los Alamos HIV Database

*“I think of it as the gift
that keeps on giving”, A. Fauci*

Quoted in Cohen, “Bang for the Buck”, Science 321:518-519, 2008

<http://tinyurl.com/HIV-DB-2018>

Los Alamos HIV Database

■ HIV Databases, funded by NIH

- Integrate HIV immunological and viral and host sequence data
- > 60 computational tools, some HIV specific; many applicable to other pathogens
- Tables, summaries, web search interfaces
- Annual Compendia

- HIV Sequence database – founded in 1986, Gerald Myers
 - Sequence data from GenBank with added metadata fields from the literature
 - Metadata and accession numbers incorporated in the sequence names
 - Premade and on the fly alignments – align indels and reduce sequences per person
 - Web searches: subtype, geographic location, patient details, sampling year, etc ~40 fields

- HIV Immunology database – founded in 1995, Bette Korber
 - Comprehensive HIV epitope database,
 - Integrates HIV immunological and sequence data
 - Web searches: epitope, protein, HLA type, immunogen, keywords, patient details, etc

■ Other pathogen databases

- HCV Database – founded in 2003, Carla Kuiken, initially funded by NIH

- HFV Database – founded in 2009, C. Kuiken, initially funded by DoD, >80 viral species
 - Filovirus portion of the database was updated during and after the 2015 outbreak
 - Premade sequence alignments on genus, species and one-per-outbreak sequence levels
 - Epitope lists and genomic maps, functional domains
 - Ebola Genome browser

HIV Database Workshop Logistics

■ Day 1, Jan 30, Tues

- HIVSequence Database

■ Day 2, Jan 31, Wed

- HIV Immunology Database

■ Part 1:

- HIV Immunology Database overview
- Antibody searches and entries in HIV database
- Neutralizing Antibody Resources
- CATNAP, both tailored for HIV and applicable to any pathogen
- CombiNaber, applicable for any pathogen
- HIV Genome Browser

■ Part 2:

- T cell epitopes and searches and entries in HIV database
- More computational tools for Immunologists, many applicable for any pathogen
- Vaccine design and evaluation tools, applicable for any pathogen



HIV DATABASES

The HIV **databases** contain data on HIV genetic sequences, immunological epitopes, drug resistance-associated mutations, and vaccine trials. The website also gives access to a large number of tools that can be used to analyze these data. This project is funded by the Division of AIDS of the National Institute of Allergy and Infectious Diseases (NIAID), a part of the National Institutes of Health (NIH). Click on any of the links below to access a database. [Editorial Board](#)

SEQUENCE DATABASE ►

VACCINE DATABASE ►

IMMUNOLOGY DATABASE ►

OTHER VIRUSES ►

News:

[Archived News ►](#)

[CATNAP: Custom Input](#)

The original CATNAP tool can compile, analyze and tally neutralizing antibody panels from a database of publicly available HIV neutralization data. A new version, [CATNAP: Custom Input](#), is now available. This version allows users to input their own neutralization panel data and perform the same analyses. HIV Env sequences are available as a premade alignment, or can be provided by the user. 12 March 2015

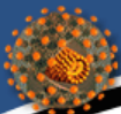
[HIV Molecular Immunology 2014](#)

HIV Molecular Immunology 2014 is now available online. The PDF version is hypertext enabled and features clickable table-of-contents, indexes, references and links to external web sites. 04 February 2015

[2014 HIV Sequence Compendium](#)

2014 was the last year that we printed and shipped the HIV Sequence Compendium. Printed copies of the 2014 compendium are still available [on request](#). 21 January 2015

www.hiv.lanl.gov



HIV sequence database

All kinds of basic information about HIV and about our database

DATABASES SEARCH ALIGNMENTS TOOLS PUBLICATIONS GUIDES Search Site

Tutorials and Basic Information

Previous workshop presentations

Tutorials

[Keystone 2014 HIV sequence database workshop](#)

[Keystone 2014 HIV Immunology database workshop](#)

[Sequence quality control](#) explains several common problems with sets of viral sequences

[How to make a phylogenetic tree](#) explains how to build a phylogenetic tree

[How to use these databases](#) summaries of workshops given at conferences

[HIV numbering](#) relative to reference strain HXB2

[SIV numbering](#) relative to reference strain SIVmm239

Articles

[3D views of HIV macromolecular structures](#) provides links to 3D views of HIV proteins

[Stalking the AIDS Virus \[PDF\]](#) article from LANL Research Quarterly (Fall 2003) about HIV Database research on the HIV-immune system interaction as a step toward an AIDS vaccine

Reference Information

[Circulating recombinant](#) details about all documented CRFs

[HIV-1 gene map](#) illustrates HXB2 breakpoints

[HXB2 annotated spreadsheet \(.xls\)](#) provides a fully-annotated sequence of HXB2 with base-by-base detail

[HIV and SIV subtype nomenclature](#) gives an overview of HIV and SIV subtype nomenclature, particularly HIV-1 groups and subtypes

[Primate immunodeficiency virus nomenclature](#) lists SIV species and nomenclature

[How the HIV database classifies sequences](#) explains how recombinants are named and annotated

[Common sequence formats for alignments](#) shows examples of common sequence formats for alignments

[How to cite this Database](#) explains how to cite this website and the printed HIV compendia

[Codes and symbols in sequences](#) decodes the symbols and IUPAC codes that appear in sequences and alignments

[Codon table](#) gives the translation of nucleotides into amino acids

[FAQs](#) answers basic questions about the HIV Sequence Database

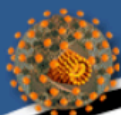
[Links](#) HIV/AIDS resources and bioinformatics tools on other websites

Yes! We do respond to this e-mail address!

last modified: Tue Aug 8 12:41 2017

Questions or comments? Contact us at seq-info@lanl.gov





HIV sequence database

DATABASES	SEARCH	ALIGNMENTS	TOOLS	PUBLICATIONS	GUIDES		Search Site
Sequence DB							
Immunology DB							
Vaccine DB							
HCV DB							
HFV DB							

HIV Sequence Database

Programs and Tools

[Search Interface](#) retrieves HIV and SIV sequences, which can then be aligned and used to build trees

[Geography Search Interface](#) retrieves HIV sequences based on geographical distribution

[Genome Browser](#) uses jBrowse to display diverse data about the HIV-1 genome and proteome

[Tools for working with sequences](#) lists all our online tools, organized by function

Alignments

[HIV Premade Alignments](#) includes Consensus and Ancestral Sequences, Subtype Reference Alignments, and Complete Alignments

Information

[HIV Sequence Compendium](#) print or order our annual publication

[Tutorials and other information](#) unpublished web-based content

[Links](#) to other HIV/AIDS tools and information

About this website

[FAQ](#) general information about this website

[Site Statistics](#) usage information for www.hiv.lanl.gov

[How to Cite this Database](#)

[Editorial Board](#)

News:

[Archived News](#) ▶

[IQ-TREE interface](#)

IQ-tree is a fast and effective stochastic algorithm for finding ML trees. We have developed a convenient web server for building trees with this method. A nice feature of this method is the ability to output a table of site-specific rates of evolution for each position in the alignment. *18 September 2017*

[IEDB User Workshop 2017](#)

The Immune Epitope Database (IEDB) will hold its 2017 User Workshop on October 25-26, 2017 in Rockville, Maryland. Staff from the LANL HIV Databases will be there to talk about our Immunology Database, Sequence Database, and bioinformatics tools. More information is available at <http://workshop.iedb.org/>. *18 July 2017*

HIV Immunology Database Overview

- Experimentally characterized immunological and associated viral data
- Key information from each paper on HIV T cell epitopes or mAbs
 - ~10,000 CTL, >1,500 Helper epitopes and >3,000 Antibody records
 - Epitope sequence, location, immunogen, vaccine details, patient details...
 - Epitope Variants (escape, reduced binding, etc.)
 - Host HLA or MHC, Ab isotype, binding region
 - Neutralizing Antibody Resources, contact residues, etc.
 - Notes summarize main findings
- HIV T cell epitopes and Antibody data organization
 - T Cells (CTL and Helper epitopes)
 - One reference per entry, epitope/HLA combinations are often repeated
 - CTL and T-helper database organization is identical
 - B Cells (Antibodies)
 - One entry for each monoclonal antibody
 - Many references per entry (> 800 for some well studied mAbs)
 - Antibody is entered and annotated whether or not epitope is defined
- HIV Immunology Database products
 - Epitope maps, summary tables and yearly compendium
 - Computational tools for immunologists
 - Neutralizing antibody resources

Tools for Immunologists

Most tools are applicable to any organism and some to any numerical data

- **CATNAP:** Compile, Analyze and Tally published and your own NAb Panels
- **CombiNAber:** Predict and analyze neutralization by antibody combinations
- **Sequence Locator:** Find epitope location on the reference genome
- **PepMap:** Map an input set of peptides on the reference sequence (Fasta, PDF and HTML)
- **PeptGen:** Generate sets of overlapping peptides for epitope mapping.
- **QuickAlign and AnalyzeAlign:** Align query sequences or discontinuous positions to an alignment, create WebLogos, calculate frequency by position, tally variants in an alignment
- **ELF:** Epitope Location Finder. Search query sequence for
 - Known epitopes from our HIV immunology databases
 - HLA binding motifs
 - [Epitopes predicted by the IEDB binding algorithm.](#)
- **N-Glycosite:** Find potential N-linked glycosylation sites in an alignment
- **Mosaic and Epigraph:** Generate candidate vaccine protein cocktails with optimized potential epitope coverage, calculate and visualize coverage
- **Heatmap:** Display and organize neutralization or other quantitative data.
- And more ...

Databases

Search

Tools

Products

Publications

Products

Epitope Maps

Epitope Tables

Epitope Alignments

T Cell Epitope Variants

Neutralizing Ab Resources & CATNAP

Data Sets: HLA Typing and Epitope Mapping

Tools & Links

Search Site

HIV Molecular Immunology Database

The HIV Molecular Immunology Database is an online collection of HIV-1 cytotoxic and helper T-cell epitopes and antibody binding sites.

Search Interfaces

- [CTL/CD8+ search](#)
- [T Helper/CD4+ search](#)
- [Antibody search](#)
- [CTL variant search](#)
- [T Helper variant search](#)
- [Search help](#)
- [Variant search help](#)

Database Products

- [All Database products and publications](#)
- [Epitope maps](#)
- [Epitope tables](#)
- [Epitope alignments](#)
- [T cell epitope variants and escape mutations](#)
- [Neutralizing antibody resources & CATNAP](#)
- [The HIV Molecular Immunology Compendium](#)
- [About the HIV Molecular Immunology Database](#)
- [How to cite this database](#)
- [Frequently-asked Questions \(FAQ\)](#)


Tools and Data Sets

- [Tools & Links](#) for immunologists
- [SIV Epitopes \(PDF\)](#) review article summarizing known SIV epitopes
- [Identifying HLA-Associated Polymorphisms in HIV-1 \(PDF\)](#) review article summarizing HIV polymorphism associated with escape mutations. Also a [table of polymorphisms](#).
- [HLATEM](#) HLA Typing and Epitope Mapping Data Sets
- [Standardized Assessments of Neutralizing Antibodies for HIV/AIDS Vaccine Development](#) Assay protocols from Duke Central Reference Laboratory

Multiple ways to database products and tools

Neutralizing Antibody Resources

<https://www.hiv.lanl.gov/content/immunology/index.html>



Neutralizing Antibody Resources

www.hiv.lanl.gov/content/immunology/neutralizing_ab_resources.html

Databases

Search

Tools

Products

Publications

Search Site

Neutralizing Antibody Resources

Tools

- [CATNAP: Compile, Analyze and Tally NAb Panels](#)
Analysis of panels of antibody data for identification of potential genetic signatures.
 - [Database CATNAP](#) analyzes published IC₅₀/IC₈₀ data for anti-HIV neutralizing antibodies.
 - [Custom CATNAP](#) analyzes any numerical data associated with a protein alignment.
 - [Hybrid CATNAP](#) analyzes your neutralization data together with published data.
- [CombiNAb](#)
Predict the neutralization of combinations of antibodies
- [HIV Genome Browser](#)
A customization of jBrowse displaying genome and proteome features of HIV, including epitopes and neutralizing antibody features.
- [External Tools for Germline Antibody Reconstruction](#)
A list of external computational tools for modeling antibody evolution and germ line reconstruction from antibody or T-cell receptor sequence data.

2 new tools useful for Ab analysis

Coming soon:

- [Genetic Signature tool](#)
 - Finds phylogenetically corrected genetic signatures in a sequence alignment in conjunction with a phenotype file.
- [Filtered Forests](#)
 - Machine learning predictions of bNAb viral sensitivity

Search interface

- [Neutralizing antibody contexts and features](#)
Search for locations of important neutralizing antibody binding sites and other HIV-1 Env features.

Tables

- [Neutralizing antibody contexts and features \(.xls\)](#)
A summary of the information from the search interface above, presented in a single .xls spreadsheet. Each row of the table corresponds to one residue of HIV-1 Env, and each column represents a protein feature or set of known binding residues of broadly neutralizing antibodies. Loops and other features of Env are shown in the first 3 columns on the left. The entropy (sequence variability) of each residue is presented numerically and color coded. Abbreviated references are listed under each column heading, and full references are on the second page of the Excel file.
- [Best neutralizing antibodies](#)
A table presenting the most broadly-neutralizing HIV-1 antibodies, with links to papers, Ab sequences, structures, notes on breadth of neutralization, locations of Ab contacts or key residues, and heavy and light chain composition.

Protocols

- [Standardized Assessments of Neutralizing Antibodies for HIV/AIDS Vaccine Development Assay protocols](#) from Duke Central Reference Laboratory

Questions or comments? Contact us at immuno@lanl.gov

Operated by Los Alamos National Security, LLC, for the U.S. Department of Energy's National Nuclear Security Administration

Copyright © 2006-2012 LANL LLC All rights reserved | [Disclaimer/Privacy](#)

HIV Molecular Immunology Database

The HIV Molecular Immunology Database is an annotated, searchable collection of HIV-1 cytotoxic and helper T-cell epitopes and antibody binding sites.

Search Interfaces

- [CTL/CD8+ search](#)
- [T Helper/CD4+ search](#)
- [Antibody search](#)
- [CTL variant search](#)
- [T Helper variant search](#)
- [Search help](#)
- [Variant search help](#)

Database Products

- [All Database products and publications](#)
- [Epitope maps](#)
- [Epitope tables](#)
- [Epitope alignments](#)
- [T cell epitope variants and escape mutations](#)
- [Neutralizing antibody resources & CATNAP](#)
- [The HIV Molecular Immunology Compendium](#)
- [About the HIV Molecular Immunology Database](#)
- [How to cite this database](#)
- [Frequently-asked Questions \(FAQ\)](#)

Tools and Data Sets

- [Tools & Links](#) for immunologists
- [SIV Epitopes \(PDF\)](#) review article summarizing known SIV epitopes
- [Identifying HLA-Associated Polymorphisms in HIV-1 \(PDF\)](#) review article summarizing HIV polymorphism associated with escape mutations. Also a [table of polymorphisms](#).
- [HLATEM](#) HLA Typing and Epitope Mapping Data Sets
- [Standardized Assessments of Neutralizing Antibodies for HIV/AIDS Vaccine Development](#) Assay protocols from Duke Central Reference Laboratory

Epitope Tables

These tables summarize the epitopes from our database. HIV-1 epitope data may also be obtained in the form of downloadable [maps](#) or [alignments](#).

- [CTL epitopes](#)
- [Best-defined \("A-list"\) CTL epitopes](#)
- [CTL epitope variants and escape mutations](#)
- [T-helper epitopes](#)
- [T Helper epitope variants and escape mutations](#)
- [Antibody epitopes](#)
- [Best Neutralizing Antibodies](#)
- [Antibody-Dependent Cell-Mediated Cytotoxicity \(ADCC\)](#)
- [Antibody index by name](#)
- [Antibody index by binding type](#)
- [SIV epitopes](#)
- [Neutralizing antibody resources](#)

Epitope alignments: epitopes aligned to HIV subtype Reference sequences in Fasta format

Reactive peptide maps and tables (with HLA and other patient data) from several large-scale studies scanning HIV proteins.

<https://www.hiv.lanl.gov/content/immunology/index.html>

Antibody Search

- An example workflow following from antibody search:
 - Search database for a particular antibody record
 - For a neutralizing antibody collect comparative neutralization data for that antibody tested against different viruses and in different studies (CATNAP)
 - Estimate the effectiveness of multi-antibody cocktails against different viruses (CombiNAber)

Antibody Search (https://www.hiv.lanl.gov/content/immunology/ab_search)

HIV protein	Proteins with defined epitopes - ALL - p17 p17-p24 p24 p24-p2p7p1p6	Proteins with undefined epitopes - ALL - p24 Gag RT Pol
HXB2 location	<input type="text"/> - <input type="text"/> Results overlap with query location	
Epitope	<input type="text"/> Results contain query sequence	
Record number	<input type="text"/>	
MAb ID	(List by name) (List by type)	
Subtype	- ALL -	
Immunogen	- ALL - anti-idiotypic autoimmune disease HIV-1 exposed seronegative HIV-1 infection HIV-2 infection in vitro stimulation or selection	
Vaccine details if Immunogen is Vaccine	Vaccine type - ALL - Vaccine strain - ALL - Vaccine component - ALL - Adjuvant - ALL -	
Ab Type	- ALL - C-domain C-HR C-term Env oligomer flap region gp120 adjacent to CD4BS	
Species	- ALL -	
Isotype	- ALL - IgA IgA1 IgA2 IgA22a IgE IgG	
Author	Search only for <input type="checkbox"/> First <input type="checkbox"/> Last author <input checked="" type="radio"/> Show only this author's references <input type="radio"/> Show all references	
Country	- ALL -	
Keywords	- ALL - acute/early infection ADCC adjuvant comparison antibody binding site definition and exposure antibody generation antibody interactions	
Note	<input checked="" type="radio"/> Show only notes containing selected keyword(s) <input type="radio"/> Show all notes <input checked="" type="radio"/> Show only notes matching this text <input type="radio"/> Show all notes	

[Click for Search Help](#)

Search by

- HIV protein, Epitope Sequence, Subtype, Immunogen, Vaccine Details, Species, Author, Country, Keywords, Isotype

■ MAb ID

- ☐ List by Ab name
- ☐ List by Ab type
 - By binding site, for example binding to similar region like V3 or near a common functional domain like CD4 binding site CD4Bs)

■ Search examples:

- ☐ 2F5 – 1 record with 815 references
- ☐ Ab type: gp120 CD4BS – 438 records

☐ Search for 10E8

Can show only notes and references containing selected keywords or user's text

Found 30 matching records:

Displaying record number 2708

Mab ID	10E8	Link to Epitope Map	gp160 Epitope Map
HXB2 Location	gp160(671-683) DNA(8235..8273)		
Author Location		Link to Epitope Alignment	Epitope Alignment
Epitope	NWFDISNWLWYIK		
Subtype	B		
Ab Type	gp41 MPER (membrane proximal external region)		
Neutralizing	P (tier 2) View neutralization details	Link to CATNAP	
Contexts and Features	Search for contexts and features	Link to Antibody Features Database (Ab contact positions and related protein features)	
Species (Isotype)	human(IgG3)		
Patient	Donor N152	Link to patient Donor detail	
Immunogen	HIV-1 infection		
Keywords	ADCC, antibody binding site, antibody gene transfer, antibody generation, antibody lineage, antibody sequence, binding affinity, bispecific molecule, broad neutralizer, chimeric antibody, computational epitope prediction, contact residues, glycosylation, immunoprophylaxis, immunotherapy, neutralization, review, structure, subtype comparisons, vaccine antigen design, vaccine-induced immune responses, variant cross-reactivity		

Notes

Showing 44 of 44 notes.

Notes from the papers

- 10E8: Next generation of a computational neutralization fingerprinting (NFP) as a way to predict polyclonal Ab responses to HIV infection is presented. A new panel of 20 pseudoviruses, termed f61, was developed to aid in the assessment of experimental neutralization. This panel was used to assess 22 well-characterized bNAbs and mixtures thereof (HJ16, VRC01, 8ANC195, IGg1b12, PGT121, PGT128, PGT135, PG9, PGT151, 35O22, 10E8, 2F5, 4E10, VRC27, VRC-CH31, VRC-PG20, PG04, VRC23, 12A12, 3BNC117, PGT145, CH01). The new algorithms accurately predicted VRC01-like and PG9-like antibody specificities. [Doria-Rose2017](#) (neutralization, computational epitope prediction)
- 10E8: The amino acid at gp120 position 375 is embedded in the Phe43 cavity, which affects susceptibility to ADCC. Most M-group strains of HIV-1 have serine at position 375, but CRF01 typically has histidine, which is a bulky residue. MAb 2G12 and 10E8 were not affected by changes in residue 375, while recognition by CD4i mAbs 17b and A32 was increased by mutations of residue 375 to histidine or tryptophan. Participants in the AIDSVAX vaccine trial were infected by CRF01, and a significant part of the efficacy of this vaccine rested on ADCC responses. The ADCC response of MAbs derived from AIDSVAX participants (CH29, CH38, CH40, CH51, CH52, CH54, CH77, CH80, CH81, CH89, CH91, CH94) was dependent on the presence of 375H and greatly decreased by the presence of 375S. [Prevost2017](#) (ADCC, vaccine-induced immune responses)

10E8 Donor

[Databases](#)[Search](#)[Tools](#)[Products](#)[Publications](#) [Search Site](#)

Patient Detail

Patient Code	Donor N152
Patient Sex	Male
Risk Factor	
Infection Country	
Infection City	
Infection Year	
HLA Type	
Patient Ethnicity	
Progression	Slow progressor (SP)
Species	human
Patient Note	At time of leukapheresis for MAb 10E8, patient had been infected with HIV-1 for 20 years; infected with clade B; selected because his serum neutralizing activity was among the most potent and broad in the cohort.
CTL CD8+ Records	
T-Helper CD4+ Records	
Antibody Records	2708 , 2709 , 3072 , 3073 , 3074 , 3075 , 3076 , 3077 , 3078 , 3079 , 3080 , 3081 , 3147 , 3148 , 3149 , 3150 , 3151 , 3152 , 3153 , 3154 , 3471 , 3472 , 3477 , 3518 , 3527
Sequence Database Patient ID Record	75177

Link to patient's HIV sequences

Neutralizing Antibody Contexts & Features

Purpose: To provide exact coordinates of known neutralizing antibody binding sites and other HIV-1 Env features. The data are also summarized in a [spreadsheet \(.xls\)](#). For details, see [Help](#).

MAB name

10-1074
10-996
10E8
12A12

Antibody class

CD4bs
CD4i
CH4bs
glycan

Env AA position

315,323

Type

antibody related feature
other Env feature

Reference

Andrabi2015
Balla-Jhagjhoorsingh2013
Bhiman2015
Blattner2014

Database ID

1
2
3
4

View Neutralizing Antibody Contexts & Features

ID 18

Description 10E8 contacts

Antibody class MPER

Reference [Huang2012a](#)

Type antibody related feature

MAB name [10E8](#) (Click MAB name to get to Immunology DB notes)

Env pos.	Feature	HXB2 AA	Entropy Group M	Entropy Subtype B	Entropy Subtype C	Annotation
671	gp41	N	0.779	0.669	0.885	10E8 N671 structure and neutralization: key epitope position.
672	gp41	W	0.017	0.023	0.014	10E8 W672 structure and neutralization: key epitope position.
673	gp41	F	0.058	0.065	0.073	10E8 F673 structure and neutralization: key epitope position.
676	gp41	T	0.683	0.610	0.674	10E8 T676 structure and binding: key epitope position.
680	gp41, gp41 transmembrane	W	0.069	0.083	0.081	10E8 W680 structure and neutralization: key epitope position.
683	gp41, gp41 transmembrane	K	0.577	0.499	0.569	10E8 K/R683 structure: key epitope position.

Important position(s) with Hxb2 amino acid: N671 W672 F673 T676 W680 K683

Submit

Reset

[Go to CATNAP main page](#)

Antibody information

Number of antibodies: 1

Download heavy and light ☒ aa ☐ na sequences in

Download table below

Expand table below to show heavy and light chain sequences and sources for germline data

Antibody	Antibody binding type	Structure	Donor	Clonal lineage	Isolation paper	Neutralizing antibody feature	Heavy V (IGHV)	Heavy D (IGHD)	Heavy J (IGHJ)	Light V (IGKV or IGLV)	Light J (IGKJ or IGLJ)	Light chain type	Genetic signature analysis	LANL comments
10E8	<ul style="list-style-type: none"> C-term gp41 MPER (membrane proximal external region) 	4U6G 5IQ7 5IQ9 4G6F	Donor N152	10E8	Huang2012a	<ul style="list-style-type: none"> 10E8 contacts 10E8 residue prediction 10E8 signature predictions (West2013) 	3-15*05	3-3*01	1*01	3-19*01	3*02	L	IC₅₀ IC₈₀	

Expand the table to show heavy and light chain sequences and sources for germline data

Link to structure in PDB

Assay

Analyze assay data in CATNAP

Number of data: 1551

Download table below with additional virus info

Expand table below to show virus information

Antibody	Virus	Reference	IC50	Mean IC50	IC80	Mean IC80
10E8	0013095_2_11	Asokan et al. J Virol 89:12501 (2015)	0.002	0.00454	0.058	0.07723
		Chuang et al. J Virol. 87:10047 (2013)	0.013			
		Doria-Rose et al. J Virol. 90:76 (2016)	0.00200		0.05800	
		Huang et al. Immunity 45:1108 (2016a) - dataset 1	0.003			
		Huang et al. Nature 491:406 (2012a)	0.003		0.069	
		Kong et al. J Virol 89:2659 (2015) - dataset 1	0.017		0.194	
		Kong et al. J Virol 89:2659 (2015) - dataset 2	0.005		0.061	

10E8 Neutralization information in CATNAP

CATNAP

Compile, Analyze and Tally NAb Panels

Purpose:

- To compile published data on HIV NAb and their neutralization data.
- To integrate and juxtapose on one screen neutralization data (or any numerical data) and viral sequence data.
- To explore potential genetic signatures associated with HIV neutralization based on either published or your own data.
- To find potential genetic signatures in any kind of numerical data associated with sequences.
- With input from Anthony West (*West et al, PNAS 2013*).
- Designed by Hyejin Yoon, Jennifer Macke, Bette Korber, Karina Yusim



CATNAP
(theoretical approximation)
Photo by Peter Hrabec

<https://www.hiv.lanl.gov/components/sequence/HIV/neutralization/index.html>

<http://hiv.lanl.gov/catnap>

CATNAP

Compile, Analyze and Tally NAb Panels

The CATNAP family of tools has been designed to facilitate the analysis of neutralizing antibodies (NAbs) through the identification of potential genetic signatures resulting from a NAb's interaction with a protein. While interactions between NAbs and HIV-1 Env are the emphasis, the Custom Input version can accommodate other types of data, including other proteins and organisms.

CATNAP

Purpose: Analyze our database of HIV-1 IC₅₀ and IC₈₀ neutralization data from publicly-available sources, in conjunction with HIV-1 Envelope sequences. Access our extensive databases of information about neutralizing antibodies and viruses used in published neutralization studies. Alignments of Env sequences for these viruses are also provided.

Help: [CATNAP Help](#).

CATNAP: Custom Input

Purpose: Find potential genetic signatures based on your own numerical data in association with protein sequences. In addition to neutralization data, this tool is flexible enough to accommodate almost any kind of data in conjunction with almost any protein sequence.

Help: [Custom CATNAP Help](#).

CATNAP: Hybrid

Purpose: Compare and analyze your HIV-1 IC₅₀ and IC₈₀ neutralization data with published data. This tool will display your data side-by-side with data from our database of published HIV-1 neutralization data.

Help: [Hybrid CATNAP Help](#).

Reference

Yoon et al. CATNAP: a tool to compile, analyze and tally neutralizing antibody panels. Nucleic Acid Res 2015 Jul 1;43(W1):W213-9. PMID 26044712.

■ Custom Input requires

- ☐ Numerical data (IC₅₀, ID₅₀, AUC, any phenotypic data)
- ☐ Aligned sequences associated with the data

■ You can also combine your own HIV data with the published HIV data (Hybrid CATNAP)

CATNAP

Compile, Analyze and Tally NAb Panels

Purpose: To provide easy analysis of data associated with HIV-1 neutralizing antibodies, including neutralization panel data, sequences, and structures.

See also: [Help](#) | [Other CATNAP tools](#) | [How to Cite](#)

[Download CATNAP data](#)

New! Click "Attributes" to select antibodies based on donor, germline genes, or binding type. Or select viruses based on tier, subtype, infection stage, or coreceptor. [Details...](#)

Select by ☒ Antibody and Virus ☐ Study

Antibodies by ☒ Names ☐ Attributes

of Abs = 304, # of Ab mixtures = 25

Select	Name	Donor	# of viruses tested
<input type="checkbox"/>	10-1074	Donor 17	420
<input type="checkbox"/>	10-1074-IgG3C	Donor 17	119
<input type="checkbox"/>	10-1074V	Donor 17	200
<input type="checkbox"/>	10-996	Donor 17	121
<input type="checkbox"/>	10E8	Donor N152	432
<input type="checkbox"/>	10E8-OfH/OfL	Donor N152	21
<input type="checkbox"/>	10E8-OfH/4fL	Donor N152	21
<input type="checkbox"/>	10E8-10fH/10fL	Donor N152	21
<input type="checkbox"/>	10E8-10fH/16fL	Donor N152	21
<input type="checkbox"/>	10E8-10fH/4fL	Donor N152	21
<input type="checkbox"/>	10E8-19fH/10fL	Donor N152	21
<input type="checkbox"/>	10E8-19fH/16fL	Donor N152	21
<input type="checkbox"/>	10E8-2fH/OfL	Donor N152	21
<input type="checkbox"/>	10E8-2fH/10fL	Donor N152	21
<input type="checkbox"/>	10E8-2fH/4fL	Donor N152	21
<input type="checkbox"/>	10E8V1.1/P140	Donor N152	118

Viruses by ☒ Names ☐ Attributes ☐ Panels

of Viruses = 1007 (781 seqs available)

Select	Name	Subtype	# of Abs tested	Seq
<input type="checkbox"/>	0013095_2_11	C	147	Yes
<input type="checkbox"/>	001428_2_42	C	146	Yes
<input type="checkbox"/>	0041_V3_C18	C	23	Yes
<input type="checkbox"/>	0077_V1_C16	C	72	Yes
<input type="checkbox"/>	00836_2_5	C	71	Yes
<input type="checkbox"/>	0260_V5_C1	A1	11	Yes
<input type="checkbox"/>	0260_V5_C36	A1	162	Yes
<input type="checkbox"/>	0301_BM_A12	C	12	Yes
<input type="checkbox"/>	0301_BM_A2	C	12	Yes
<input type="checkbox"/>	0301_BM_A6	C	12	Yes
<input type="checkbox"/>	0330_V4_C3	A1	121	Yes
<input type="checkbox"/>	0404_BM_B9	C	12	Yes
<input type="checkbox"/>	0404_BM_D4	C	12	Yes
<input type="checkbox"/>	0404_BM_F3	C	12	Yes
<input type="checkbox"/>	0404_BM_G3	C	12	Yes
<input type="checkbox"/>	0404_BM_H4	C	12	Yes

Options

Retrieve ☐ Antibody details ☐ Virus details ☐ Assay

Or

Analyze along with virus sequences ☒ IC₅₀ ☐ IC₈₀ ☐ Both

Large sets of data run slowly. Limit the number of antibodies or viruses for quicker response.

☐ Exclude viruses having no sequence data

☐ Email results

Submit

Reset

Both DATABASE and ANALYSIS:

- Database of all the published IC₅₀ and IC₈₀ assays we can find (110 currently)
- Data:
 - Antibody data (>300): donor ID, links to Immuno DB, PDB structures, germline, binding type, etc.
 - Aligned virus data (>1000): subtype, accession, neutralization tier, virus name aliases, patient health status, various viral panels, etc.
 - Information about Env positions: entropy, functional domain, Ab contacts and signature predictions
- Analysis:
 - Env sequence data side-by-side with IC₅₀/IC₈₀ values
 - AA composition, N-glycosylation sites, basic statistics
 - Antibody potency and breadth summarized over multiple studies
 - Amino acid associations with neutralization.
 - Links to other analysis tools

CATNAP

Compile, Analyze and Tally NAb Panels

Purpose: To provide easy analysis of data associated with HIV-1 neutralizing antibodies, including neutralization panel data, sequences, and structures.

See also: [Help](#) | [Other CATNAP tools](#) | [How to Cite](#)

[Download CATNAP data](#)

New! Click "Attributes" to select antibodies based on donor, germline genes, or binding type. Or select viruses based on tier, subtype, infection stage, or coreceptor. [Details...](#)

Select Antibodies and Viruses in Several Ways:

- Individual or all antibody and viruses

Select by **Antibody and Virus** **Study**

Antibodies by ☒ Names ☐ Attributes

of Abs = 304, # of Ab mixtures = 25

Select	Name	Donor	# of viruses tested
<input type="checkbox"/>	10-1074	Donor 17	420
<input type="checkbox"/>	10-1074-IgG3C	Donor 17	119
<input type="checkbox"/>	10-1074V	Donor 17	200
<input type="checkbox"/>	10-996	Donor 17	121
<input type="checkbox"/>	10E8	Donor N152	432
<input type="checkbox"/>	10E8-OfH/OfL	Donor N152	21
<input type="checkbox"/>	10E8-OfH/4fL	Donor N152	21
<input type="checkbox"/>	10E8-10fH/10fL	Donor N152	21
<input type="checkbox"/>	10E8-10fH/16fL	Donor N152	21
<input type="checkbox"/>	10E8-10fH/4fL	Donor N152	21
<input type="checkbox"/>	10E8-19fH/10fL	Donor N152	21
<input type="checkbox"/>	10E8-19fH/16fL	Donor N152	21
<input type="checkbox"/>	10E8-2fH/OfL	Donor N152	21
<input type="checkbox"/>	10E8-2fH/10fL	Donor N152	21
<input type="checkbox"/>	10E8-2fH/4fL	Donor N152	21
<input type="checkbox"/>	10E8V1.1/P140	Donor N152	118

Viruses by ☒ Names ☐ Attributes ☐ Panels

of Viruses = 1007 (781 seqs available)

Select	Name	Subtype	# of Abs tested	Seq
<input type="checkbox"/>	0013095_2_11	C	147	Yes
<input type="checkbox"/>	001428_2_42	C	146	Yes
<input type="checkbox"/>	0041_V3_C18	C	23	Yes
<input type="checkbox"/>	0077_V1_C16	C	72	Yes
<input type="checkbox"/>	00836_2_5	C	71	Yes
<input type="checkbox"/>	0260_V5_C1	A1	11	Yes
<input type="checkbox"/>	0260_V5_C36	A1	162	Yes
<input type="checkbox"/>	0301_BM_A12	C	12	Yes
<input type="checkbox"/>	0301_BM_A2	C	12	Yes
<input type="checkbox"/>	0301_BM_A6	C	12	Yes
<input type="checkbox"/>	0330_V4_C3	A1	121	Yes
<input type="checkbox"/>	0404_BM_B9	C	12	Yes
<input type="checkbox"/>	0404_BM_D4	C	12	Yes
<input type="checkbox"/>	0404_BM_F3	C	12	Yes
<input type="checkbox"/>	0404_BM_G3	C	12	Yes
<input type="checkbox"/>	0404_BM_H4	C	12	Yes

Options

Retrieve ☐ Antibody details ☐ Virus details ☐ Assay

Or

Analyze along with virus sequences ☒ IC₅₀ ☐ IC₈₀ ☐ Both

Large sets of data run slowly. Limit the number of antibodies or viruses for quicker response.

☐ Exclude viruses having no sequence data

☐ Email results

CATNAP

Compile, Analyze and Tally NAB Panels

Purpose: To provide easy analysis of data associated with HIV-1 neutralizing antibodies, including neutralization panel data, sequences, and structures.

See also: [Help](#) | [Other CATNAP tools](#) | [How to Cite](#)

[Download CATNAP data](#)

New! Click "Attributes" to select antibodies based on donor, germline genes, or binding type. Or select viruses based on tier, subtype, infection stage, or coreceptor. [Details...](#)

Select by ☐ Antibody and Virus ☒ Study

of Studies = 110

Select	Name	# of Abs tested	# of viruses tested	IC ₅₀	IC ₈₀
<input type="radio"/>	Acharya et al. J Virol 87:10173 (2013)	6	195	V	
<input type="radio"/>	Andrabi et al. Immunity 43:959 (2015)	23	21	V	
<input type="radio"/>	Andrabi et al. Virology 439:81 (2013)	18	21	V	
<input type="radio"/>	Asokan et al. J Virol 89:12501 (2015)	5	33	V	V
<input type="radio"/>	Balla-Jhagjhoorsingh et al. PLoS One 6:e25488 (2011)	3	14	V	
<input type="radio"/>	Balla-Jhagjhoorsingh et al. PLoS One 8:e68863 (2013)	6	6	V	
<input type="radio"/>	Bhiman et al. Nat Med 21:1332 (2015)	30	5	V	
<input type="radio"/>	Bonsignori et al. Cell 165:449 (2016)	4	201	V	
<input type="radio"/>	Bonsignori et al. J Clin Invest 124:1835 (2014)	1	42	V	
<input type="radio"/>	Bonsignori et al. J Virol 85:9998 (2011)	5	90	V	V
<input type="radio"/>	Bonsignori et al. J Virol 86:4688 (2012)	2	97	V	
<input type="radio"/>	Bonsignori et al. Sci Transl Med 9:eaai7514 (2017)	14	209	V	
<input type="radio"/>	Bournazos et al. Cell 165:1609 (2016) - dataset 1	22	123	V	V
<input type="radio"/>	Bournazos et al. Cell 165:1609 (2016) - dataset 2	18	7	V	V
<input type="radio"/>	Bradley et al. EBioMedicine 12:196 (2016a)	15	6	V	
<input type="radio"/>	Braibant et al. AIDS 27:1239-44 (2013)	9	15	V	
<input type="radio"/>	Chaillon et al. J Virol 86:10540 (2012)	6	10	V	
<input type="radio"/>	Changela et al. J Virol. 85:2524 (2011)	2	4	V	

Options

Retrieve ☐ Antibody details ☐ Virus details ☐ Assay

Or

Analyze along with virus sequences ☒ IC₅₀ ☐ IC₈₀ ☐ Both

Large sets of data run slowly. Limit the number of antibodies or viruses for quicker response.

☐ Exclude viruses having no sequence data

☐ Email results

Select Antibodies and Viruses in Several Ways:

- Individual or all antibody and viruses
- Select by study

CATNAP

Compile, Analyze and Tally NAb Panels

Purpose: To provide easy analysis of data associated with HIV-1 neutralizing antibodies, including neutralization panel data, sequences, and structures.

See also: [Help](#) | [Other CATNAP tools](#) | [How to Cite](#)

[Download CATNAP data](#)

New! Click "Attributes" to select antibodies based on donor, germline genes, or binding type. Or select viruses based on tier, subtype, infection stage, or coreceptor. [Details...](#)

Select by **Antibody and Virus** **Study**

Antibodies by ☐ Names ☒ Attributes

(# of Antibodies) Reset

Donor	Light V (IG)	Light J (IG)
127/C (2)	KV1-1 (1)	KJ1 (14)
44 (1)	KV1-13*02 (1)	KJ1*01 (20)
C38 (3)	KV1-17*01 (1)	KJ2 (2)
CAP206 (1)	KV1-33*01 (11)	KJ2*01 (15)

Heavy V (IGHV)	Heavy D (IGHD)	Heavy J (IGHJ)
1-02*02 (11)	10 (1)	1 (2)
1-03*01 (1)	1-26 (1)	1*01 (10)
1-18*01 (2)	16 (1)	2 (9)
1-18*02 (8)	1-IR1 (1)	2*01 (5)

AB binding type

C-term (1)
Env oligomer (7)
gp120 (2)
gp120 adjacent to CD4BS (1)

Display a record if

☒ ALL selected conditions are true (intersection)
☐ AT LEAST ONE selected condition is true (union)

Viruses by ☐ Names ☒ Attributes ☐ Panels

(# of Viruses) Reset

Tier	Subtype
1 (2)	01_AE (67)
1A (6)	02A1 (5)
1A or 1B (1)	02_AG (22)
1B (23)	06_cpx (1)
1B or 2 (16)	07_BC (48)
1 or 2 (3)	08_BC (9)

Infection stage	Coreceptor
acute (171)	CCR5 (155)
intermediate (97)	CCR5 CXCR4 (6)
early (303)	CXCR4 (19)
chronic (292)	null (827)
null (142)	

Display a record if

☒ ALL selected feature conditions are true (intersection)
☐ AT LEAST ONE selected feature condition is true (union)

Options

Retrieve ☐ Antibody details ☐ Virus details ☐ Assay

Or

Analyze along with virus sequences ☒ IC₅₀ ☐ IC₈₀ ☐ Both

Large sets of data run slowly. Limit the number of antibodies or viruses for quicker response.

☐ Exclude viruses having no sequence data

☐ Email results

Select Antibodies and Viruses in Several Ways:

- Individual or all antibody and viruses
- Select by study
- Select antibodies by attributes (germline and binding region)
- Select viruses by attributes (Tier, Subtype, Infection stage)

CATNAP

Compile, Analyze and Tally Nab Panels

Purpose: To provide easy analysis of data associated with HIV-1 neutralizing antibodies, including neutralization panel data, sequences, and structures.

See also: [Help](#) | [Other CATNAP tools](#) | [How to Cite](#)

[Download CATNAP data](#)

New! Click "Attributes" to select antibodies based on donor, germline genes, or binding type. Or select viruses based on tier, subtype, infection stage, or coreceptor. [Details...](#)

Select by **Antibody and Virus** **Study**

Antibodies by ☐ Names ☒ Attributes

(# of Antibodies)

Donor	Light V (IG)	Light J (IG)
127/C (2)	KV1-1 (1)	KJ1 (14)
44 (1)	KV1-13*02 (1)	KJ1*01 (20)
C38 (3)	KV1-17*01 (1)	KJ2 (2)
CAP206 (1)	KV1-33*01 (11)	KJ2*01 (15)

Heavy V (IGHV)	Heavy D (IGHD)	Heavy J (IGHJ)
1-02*02 (11)	10 (1)	1 (2)
1-03*01 (1)	1-26 (1)	1*01 (10)
1-18*01 (2)	16 (1)	2 (9)
1-18*02 (8)	1-IR1 (1)	2*01 (5)

AB binding type

C-term (1)
Env oligomer (7)
gp120 (2)
gp120 adjacent to CD4BS (1)

Display a record if

☒ ALL selected conditions are true (intersection)
☐ AT LEAST ONE selected condition is true (union)

Viruses by ☐ Names ☐ Attributes ☒ Panels

of Panels = 3

Select	Name	Reference	# of viruses
<input type="checkbox"/>	C clade	Rademeyer2016	200
<input type="checkbox"/>	Global	Decamp2014	12
<input type="checkbox"/>	Most common		200

Options

Retrieve ☐ Antibody details ☐ Virus details ☐ Assay

Or

Analyze along with virus sequences ☒ IC₅₀ ☐ IC₈₀ ☐ Both

Large sets of data run slowly. Limit the number of antibodies or viruses for quicker response.

☐ Exclude viruses having no sequence data

☐ Email results

Select Antibodies and Viruses in Several Ways:

- Individual or all antibody and viruses
- Select by study
- Select antibodies by attributes (germline and binding region)
- Select viruses by attributes (Tier, Subtype, Infection stage)
- Select viruses by a virus panel

CATNAP

Compile, Analyze and Tally NAb Panels

Purpose: To provide easy analysis of data associated with HIV-1 neutralizing antibodies, including neutralization panel data, sequences, and structures.

See also: [Help](#) | [Other CATNAP tools](#) | [How to Cite](#)

[Download CATNAP data](#)

New! Click "Attributes" to select antibodies based on donor, germline genes, or binding type. Or select viruses based on tier, subtype, infection stage, or coreceptor. [Details...](#)

Select by **Antibody and Virus** **Study** ⓘ

Antibodies by ☒ Names ☐ Attributes ⓘ

of Abs = 307, # of Ab mixtures = 25

Select	Name	Donor	# of viruses tested
<input type="checkbox"/>	10-1074	Donor 17	420
<input type="checkbox"/>	10-1074-IgG3C	Donor 17	119
<input type="checkbox"/>	10-1074V	Donor 17	200
<input type="checkbox"/>	10-996	Donor 17	121
<input checked="" type="checkbox"/>	10E8	Donor N152	433
<input type="checkbox"/>	10E8-OfH/OfL	Donor N152	21
<input type="checkbox"/>	10E8-OfH/4fL	Donor N152	21
<input type="checkbox"/>	10E8-10fH/10fL	Donor N152	21
<input type="checkbox"/>	10E8-10fH/16fL	Donor N152	21
<input type="checkbox"/>	10E8-10fH/4fL	Donor N152	21
<input type="checkbox"/>	10E8-19fH/10fL	Donor N152	21
<input type="checkbox"/>	10E8-19fH/16fL	Donor N152	21
<input type="checkbox"/>	10E8-2fH/OfL	Donor N152	21
<input type="checkbox"/>	10E8-2fH/10fL	Donor N152	21
<input type="checkbox"/>	10E8-2fH/4fL	Donor N152	21
<input type="checkbox"/>	10E8V1.1/P140	Donor N152	118

Viruses by ☒ Names ☐ Attributes ⓘ ☐ Panels ⓘ

of Viruses = 1011 (785 seqs available)

Select	Name	Subtype	# of Abs tested	Seq
<input type="checkbox"/>	0013095_2_11	C	147	Yes
<input type="checkbox"/>	001428_2_42	C	146	Yes
<input type="checkbox"/>	0041_V3_C18	C	23	Yes
<input type="checkbox"/>	0077_V1_C16	C	72	Yes
<input type="checkbox"/>	00836_2_5	C	71	Yes
<input type="checkbox"/>	0260_V5_C1	A1	11	Yes
<input type="checkbox"/>	0260_V5_C36	A1	163	Yes
<input type="checkbox"/>	0301_BM_A12	C	12	Yes
<input type="checkbox"/>	0301_BM_A2	C	12	Yes
<input type="checkbox"/>	0301_BM_A6	C	12	Yes
<input type="checkbox"/>	0330_V4_C3	A1	122	Yes
<input type="checkbox"/>	0404_BM_B9	C	12	Yes
<input type="checkbox"/>	0404_BM_D4	C	12	Yes
<input type="checkbox"/>	0404_BM_F3	C	12	Yes
<input type="checkbox"/>	0404_BM_G3	C	12	Yes
<input type="checkbox"/>	0404_BM_H4	C	12	Yes

Example:
10E8 and PG9

Retrieve Antibody, Virus or Assay details

Options

Retrieve ☒ Antibody details ☐ Virus details ☐ Assay

Or

Analyze along with virus sequences ☐ IC₅₀ ☐ IC₈₀ ☐ Both

Large sets of data run slowly. Limit the number of antibodies or viruses for quicker response.

☐ Exclude viruses having no sequence data

☐ Email results

Analyze IC₅₀, IC₈₀ or Both along with the viral sequences

[Go to CATNAP main page](#)

Antibody information

Number of antibodies: 1

Download heavy and light ☒ aa ☐ na sequences in

Download table below

Expand table below to show heavy and light chain sequences and sources for germline data

Antibody	Antibody binding type	Structure	Donor	Clonal lineage	Isolation paper	Neutralizing antibody feature	Heavy V (IGHV)	Heavy D (IGHD)	Heavy J (IGHJ)	Light V (IGKV or IGLV)	Light J (IGKJ or IGLJ)	Light chain type	Genetic signature analysis	LANL comments
10E8	<ul style="list-style-type: none"> C-term gp41 MPER (membrane proximal external region) 	4U6G 5IQ7 5IQ9 4G6F	Donor N152	10E8	Huang2012a	<ul style="list-style-type: none"> 10E8 contacts 10E8 residue prediction 10E8 signature predictions (West2013) 	3-15*05	3-3*01	1*01	3-19*01	3*02	L	IC₅₀ IC₈₀	

Expand the table to show heavy and light chain sequences and sources for germline data

Link to structure in PDB

Assay

Analyze assay data in CATNAP

Number of data: 1551

Download table below with additional virus info

Expand table below to show virus information

Antibody	Virus	Reference	IC50	Mean IC50	IC80	Mean IC80
10E8	0013095_2_11	Asokan et al. J Virol 89:12501 (2015)	0.002	0.00454	0.058	0.07723
		Chuang et al. J Virol. 87:10047 (2013)	0.013			
		Doria-Rose et al. J Virol. 90:76 (2016)	0.00200		0.05800	
		Huang et al. Immunity 45:1108 (2016a) - dataset 1	0.003			
		Huang et al. Nature 491:406 (2012a)	0.003		0.069	
		Kong et al. J Virol 89:2659 (2015) - dataset 1	0.017		0.194	
		Kong et al. J Virol 89:2659 (2015) - dataset 2	0.005		0.061	

10E8 Neutralization information in CATNAP

Virus information

Number of viruses: 14

CATNAP: Virus info (in addition to the Ab and assay info)

Automatically submit all selected sequences in a batch to the HIV sequence search interface

[Download](#) table below

More info in HIV Sequence DB

Virus name	Subtype	Country	Patient health	Days post infection	Days from seroconversion	Fiebig	Risk factor	Accession	Tier	Alias	HIV DB name	Seq data	LANL comments
216_F2_E3_5	A1C	TANZANIA				6	Heterosexual	HM215277			216_F2_E3_5	Yes	
231965_C1	D	UGANDA	acute infection		early	1 or 2		JQ361079	2	231965, 231965_C01	231965_c01	Yes	
231966_C2	D	UGANDA	acute infection		early	1 or 2		JX512899	2	231966_C02	231966_c02	Yes	
234_F1_16_57	C	TANZANIA			early	5	Heterosexual	HM215278			234_F1_16_57	Yes	
235_47	02_AG	CAMEROON				6	Not Recorded	EU513195	2	235	235	Yes	Sequence does not match accession. This sequence/clone was the one used in neutralization studies but it has not yet been deposited in GenBank.
242_14	02A1	CAMEROON				6		EU513188	1B or 2	242	242	Yes	
246_F3_C10_2	AC	TANZANIA				6	Heterosexual	HM215279			246_F3_C10_2	Yes	
246F_C1G	C	ZAMBIA	acute infection		early	2	Heterosexual	FJ496194	2	ZM246, 246F	ZM246F_f1D5	Yes	
247_23	DU	CAMEROON					Not Recorded	EU683891	2	247	247	Yes	
249M_B10	C	ZAMBIA	acute infection		early		Heterosexual	EU166862	2	249M	ZM249M_080503_SGA_B10	Yes	
25710_2_43	C	INDIA	acute infection		45	5	Heterosexual	EF117271	1B or 2	25710	HIV_25710_2	Yes	
25711_2_4	C	INDIA	acute infection		45	3	Heterosexual	EF117272	1B or 2	25711	HIV_25711_2	Yes	
25925_2_22	C	INDIA	acute infection		45	3	Heterosexual	EF117273	1B or 2	25925	HIV_25925_2	Yes	
26191_2_48	C	INDIA	acute infection		45	3	Heterosexual	EF117274	2	26191	HIV_26191_2	Yes	

Link to the sequence record in the HIV Sequence DB

CATNAP: IC₅₀ & IC₈₀/HIV-1 alignment

Collapse or expand details from individual studies

4 antibodies & 1007 viruses selected to search

[More virus info in HIV Seq DB](#)

Virus name	Tier
001428_2_42	2
0041_V3_C18	2
0077_V1_C16	2
00836_2_5	1B or 2
0260_V5_C36	
0301_BM_A12	
0301_BM_A2	
0301_BM_A6	
0330_V4_C3	2
0404_BM_B9	
0404_BM_D4	
0404_BM_F3	
0404_BM_G3	
0404_BM_H4	
0439_V5_C1	2
0702_BM_B4	
0702_BM_B9	
0702_BM_D1	
0702_BM_H12	
0815_V3_C3	2
0907_V4_C12	
0921_V3_C14	

*: Geometric means
(Expand to individual values)

	10E8 IC50	10E8 IC80	PG9 IC50	PG9 IC80	PG1
001428_2_42	1.48106*	6.44387*	0.01010*	0.01221*	0.0
0041_V3_C18	0.44000	3.70700	0.00100	0.00400	1.8
0077_V1_C16	1.45938*	9.19287*	0.09106*	0.35301*	UD::
00836_2_5	0.50193*	1.90470*	9.00000*	UD:>50*	31.
0260_V5_C36	9.80180*	30.00417*	1.57709*	13.80908*	0.0
0301_BM_A12	-	-	0.66000	-	-
0301_BM_A2	-	-	0.91000	-	-
0301_BM_A6	-	-	0.91000	-	-
0330_V4_C3	1.04674*	3.83465*	0.01468*	0.05294*	0.0
0404_BM_B9	-	-	0.08000	-	-
0404_BM_D4	-	-	6.87000	-	-
0404_BM_F3	-	-	0.04000	-	-
0404_BM_G3	-	-	0.58000	-	-
0404_BM_H4	-	-	0.01000	-	-
0439_V5_C1	1.16920*	4.69871*	UD:>50*	UD:>50*	UD::
0702_BM_B4	-	-	0.83000	-	-
0702_BM_B9	-	-	0.38000	-	-
0702_BM_D1	-	-	0.18000	-	-
0702_BM_H12	-	-	0.37000	-	-
0815_V3_C3	0.31354*	1.86820*	UD:>10, >25, ...	*UD:>25, >50*	0.0
0907_V4_C12	-	-	-	-	-
0921_V3_C14	-	-	-	-	-
Geometric mean of detected	0.35352	2.15327	0.18501	0.38293	0.3
Geometric mean of all (undetected set to 100)	0.46515	2.43935	0.92140	2.20904	0.9
% detected (detected/total)	95% (411/432)	97% (387/400)	74% (543/729)	69% (285/416)	64%

HXB2

MRVKE---KY-QHLW-RWG---WRWGTMLLG---MLMI---CSAT---
----- ----- ----- ----- ----- -----
-----10-----20-----30-----
MRVRGILR-NY-QQW-----WMWGVLFWF---MLMI---CNGV---
MRVRGILR-NW-QLW-----WTWGILGFW---MVMN---CNVR---
MRVMGSMR-NC-QRW-----WIWGILGFW---MLMT---CNME---
MRVRGIRR-NY-QHW-----WIWGILGFW---MLMI---CKGR---
MRVMGIQR-NS-QCF-----LSWGMVLVG---IMMI---CSAV---
MRVRGMMR-NW-QQW-----WIWGILGFW---MLMI---CSVL---
MRVRGMMR-NW-QQW-----WIWGILGFW---MLMM---CSVL---
MRVRGMMR-NW-QQW-----WIWGILGFW---MLMM---CSVL---
MRVMGMQR-NS-RHL-----LLRWGIRILG---MIMI---CRTA---
MRVRGILR-NC-PQW-----WTWGILGFW---MLMI---CSVW---
MRVRGILR-NC-PQW-----WTWGILGFW---MLMI---CSVW---
MRVRGILR-NC-PQW-----WTWGILGFW---MLMI---CSVW---
MRVRGILR-NC-PQW-----WTWGILGFW---MLMI---CSVW---
MRVRGILR-NC-PQW-----WTWGILGFW---MLMI---YSVW---
MRVMGIQR-NC-QHL-----LRWGTLLG---LIII---CSTA---
MRVRGILR-NW-ELW-----WIWGILGFW---MFMI---CNML---
MRVRGILK-NW-KLW-----WIWGILGFW---MFMI---CNML---
MRVRGILR-NW-ELW-----WIWGILGFW---MFMI---CNML---
MRVRGILR-NW-KLW-----WIWGILGFW---IFMI---CNTL---
MRVMGIQM-NW-QQW-----WIWGILGFW---MLMV---CNGT---
MRAREMKR-NC-QNL-----WKWGMILLG---ILMI---CSAA---

Potency and Breadth of neutralization over multiple studies

of viruses found: 799
of Abs found: 4
of studies found: 70

[Download](#) neutralization data with virus information

[Download](#) ☒ aa ☐ na in [Fasta](#)

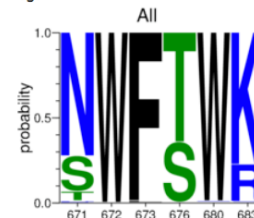
[Go to antibody information section](#)

Antibody context and feature position(s) (based on HXB2)

(See [Spreadsheet of neutralizing antibody contexts and features](#) (.xls) for more information)

- 10E8 contacts ([LogoByAll](#) [LogoBySubtype](#)): N671 W672 F673 T676 W680 K683
- PG9-like contacts ([LogoByAll](#) [LogoBySubtype](#)): N156 N160 I165 G167 K168 V169 Q170 K171 Y173

Logo



Position analysis

[Analyze](#) HXB2 position for Ab [Pick Ab and click on contact position to analyze, or enter your own position](#)

[Run CombiNaber](#) [Submit](#)

# of viruses tested						
10E8 IC50: 432	10E8 IC80: 400	PG9 IC50: 729	PG9 IC80: 634	PGT121 IC50: 634	PGT121 IC80: 393	VRC01 IC50: 781
388 virus(es) tested against all antibodies retrieved will be submitted to CombiNaber.						

Analysis at position 160 for Ab PG9

Amino Acid Counts					
AA	Count	# for detected	# for undetected	Fisher test p-value	Odds ratio
N	544	425	119	< 2.2e-16	25.55874
D	10	0	10	1.37e-06	0
K	9	0	9	5.403e-06	0
S	5	1	4	0.01897	0.0884202
Y	5	1	4	0.01897	0.0884202
X	4	3	1	1	1.081824
R	3	0	3	0.01834	0
T	1	0	1	0.265	0
V	1	0	1	0.265	0
H	1	0	1	0.265	0
-	1	0	1	0.265	0
I	1	0	1	0.265	0
Total	585	430	155		
no seq	144				
Grand total	729				

N-linked Glycosylation Motif Counts					
NxST	Count	# for detected	# for undetected	Fisher test p-value	Odds ratio
g+	531	424	107	< 2.2e-16	31.48806
g-	53	6	47	< 2.2e-16	0.03273309
-	1	0	1	0.265	0
Total	585	430	155		
no seq	144				
Grand total	729				

Note: The new Genetic Signature Tool calculating phylogenetically corrected signatures will be linked soon to CATNAP (pending submitted publication)

Odds ratio > 1: enriched for detected
Odds ratio < 1: enriched for undetected

HXB2	
IEKGEIKNCSF	NISTSIRG-KVQKEYAFFYKLDIIPIDN-----DTT
----- ----- ----- -----	
0-----160-----170-----180-----19	
AA (NxST)	
N (+)	YKEDIRNCSF
N (+)	NGDEMKNCSF
N (+)	TSNEMKNCSF
N (+)	YESMKNCSF
N (+)	MEGEIKDCSF
N (+)	TRDELNCSY
N (+)	TENERKNCSF
R	ISTADMKNCSF
N (+)	IMTNCTFN
N (+)	DKGEMKKNCSF
N (+)	ESGEIKNCSF
N	IDPGEIKNCSF
N (+)	IEKGEIKNCSF
N (+)	NGEEIKNCSF
N (+)	DMGEIKNCSF
N (+)	INVEEMKNCSF
N (+)	MEGEIKNCSF
N (+)	PEAGMKNCSF
N (+)	IQGEEMKNCSF
N (+)	IMKGEITNCSF
N (+)	INTEDMKNCSF

About this position
Position: Env 160 (193 in alignment above)
Entropy, M group: 0.401
Functional domain: gp120 (Kwong2000), V2 (Leonard1990)
Antibody features of this position
Mutation affects PG9-like Ab sensitivity: Loss of glycan confers resistance; PG9-like class includes PG16, PGT141, 145, CH01-CH04 (V1V2 glycan, Doria-RoseNA2012)
PG16 signature predictions: PG16: glycosylation at N160 is associated with increased susceptibility to neutralization; intermediate quality of support. (V1V2 glycan, West2013)
PG9-like contacts: PG9 glycan contact; PG9-like class includes PG16, PGT141, 145, CH01-CH04 (V1V2 glycan, McLellan2011)
PG9 signature predictions: PG9: N160 is associated with increased susceptibility to neutralization; intermediate quality of support. (V1V2 glycan, West2013)

(For more information, check Neutralizing Antibody Contexts & Features)

CombiNAber

A tool for Prediction & Analysis of Neutralization by Antibody Combinations

Purpose: This tool predicts and analyzes combination antibody neutralization scores using IC₅₀ and/or IC₈₀ for individual antibodies. The predicted scores are systematically compared for all single antibodies and 2, 3 and 4 antibody combinations analyzed. See [explanation](#).

IC₅₀/IC₈₀ data

Paste values or upload file

(See [assay requirements](#)) '<' and '>' signs are NOT allowed. Please replace them with 'LT' and 'GT' respectively.

[Sample Input]

No file selected.

Data type ☐ IC₅₀ ☐ IC₈₀ ☒ Both

Delimiter ☐ Comma ☐ Space ☒ Tab

mAb class

Paste values or upload file

(See [Ab class requirements](#))

No file selected.

Delimiter ☐ Comma ☐ Space ☒ Tab

Options

Prediction method ☒ Additive ☐ Bliss-Hill

mAb combinations ☒ Combinations using full set of mAbs

of Abs in Ab combination ☒ 2 ☒ 3 ☒ 4 (may be adjusted depending # of Abs)

☐ Repeat mAbs from same class in combinations

Combinations of interest ([example](#))

No file selected.

Analyses Target concentration ug/ml (separate with commas if more than one concentration)

Active coverage by multiple mAbs in combination ☐ 2 ☐ 3 ☒ 4

☐ Incomplete neutralization

☐ Instantaneous inhibitory potential (IIP)

File format for figures ☐ PDF ☐ SVG ☒ PNG

Email results ☐

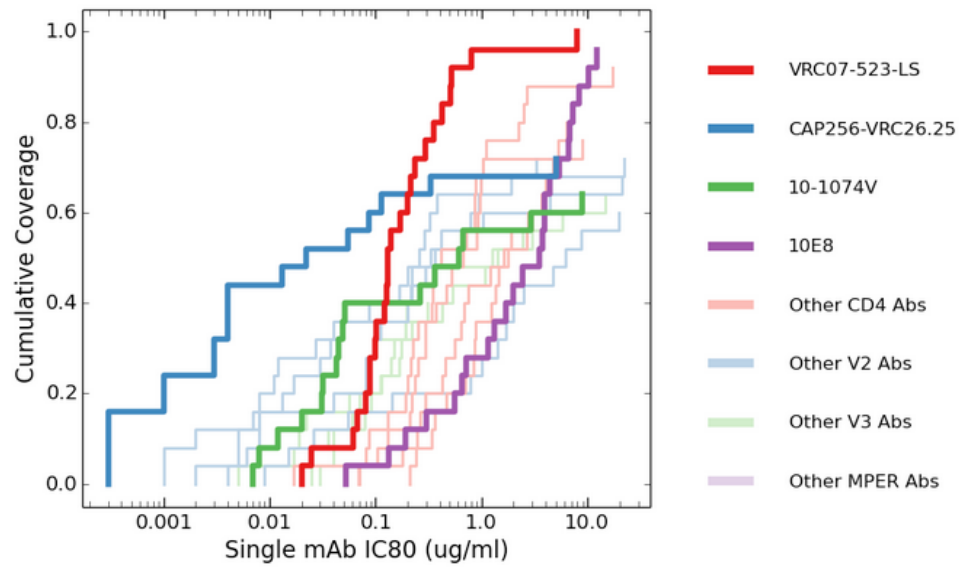
CombiNAber

- Our newest tool, designed by [Kshitij Wagh](#), [Hyejin Yoon](#), [Bette Korber](#)
- Background
 - Kong *et al*, 2015, *J Virol*
 - Wagh *et al*, 2016, *PLOS Pathogens*
 - Questions: [Kshitij Wagh](#), kshitij@lanl.gov
- Purpose: predict neutralization by antibody combinations (to optimize immunotherapy options)
- Input:
 - Neutralization data (IC₅₀ and / or IC₈₀) with antibody and virus names
 - Antibody type (i.e. binding region)

www.hiv.lanl.gov/content/sequence/COMBINABER/combinaber.html

CombiNAber

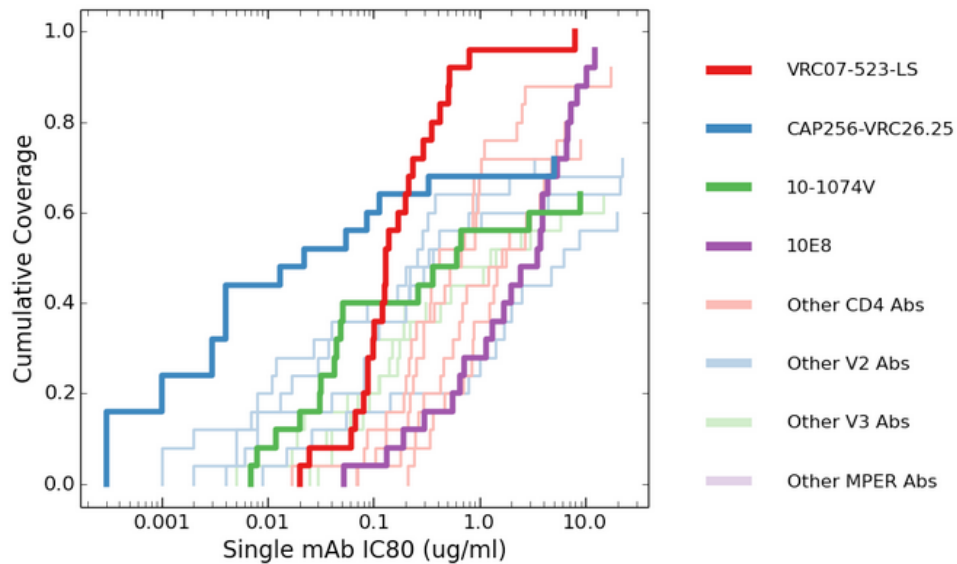
Overall breadth potency ?



Single mAbs

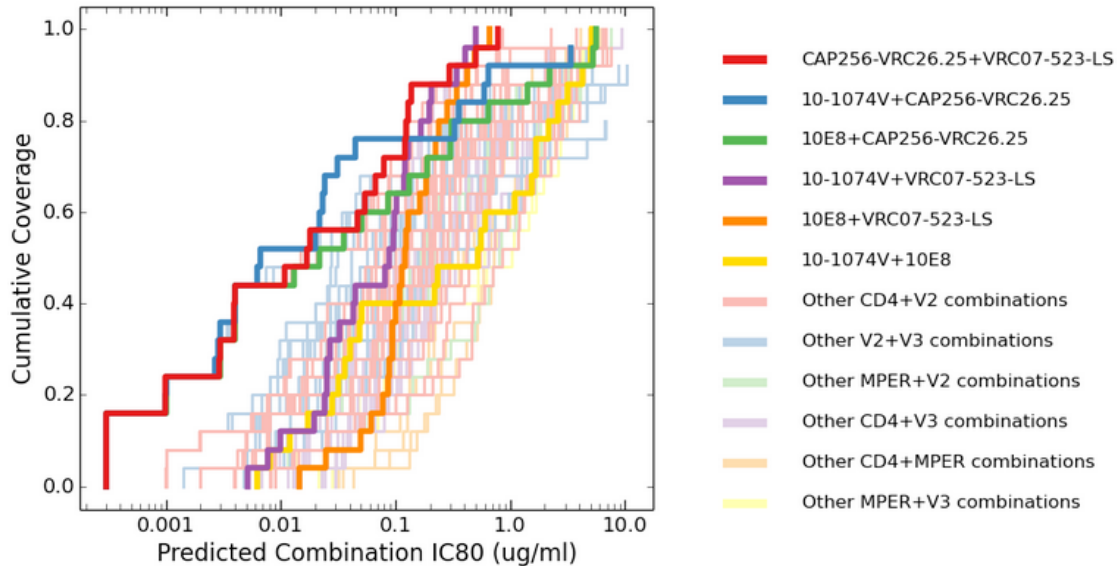
CombiNAber

Overall breadth potency ?



Single mAbs

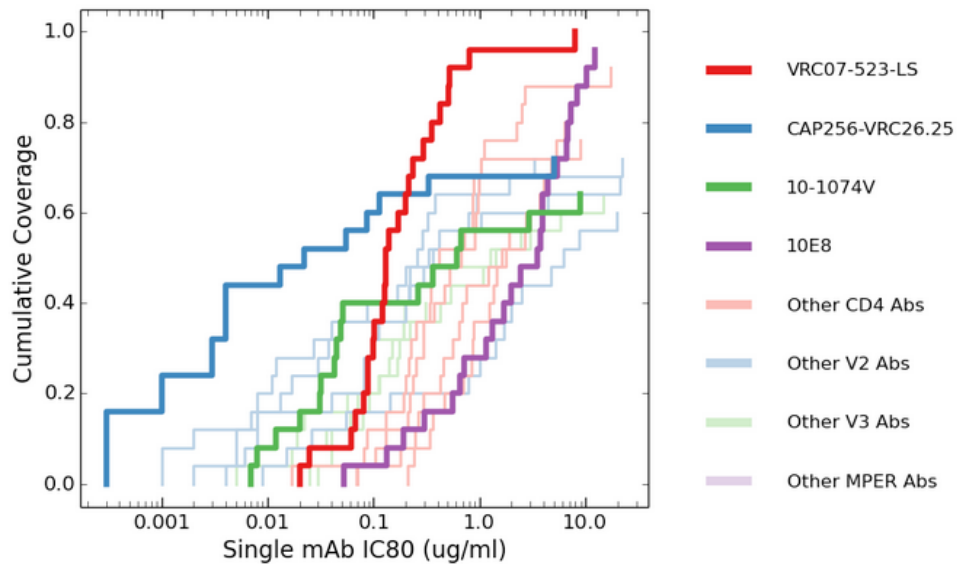
Overall breadth potency ?



2-mAb combinations

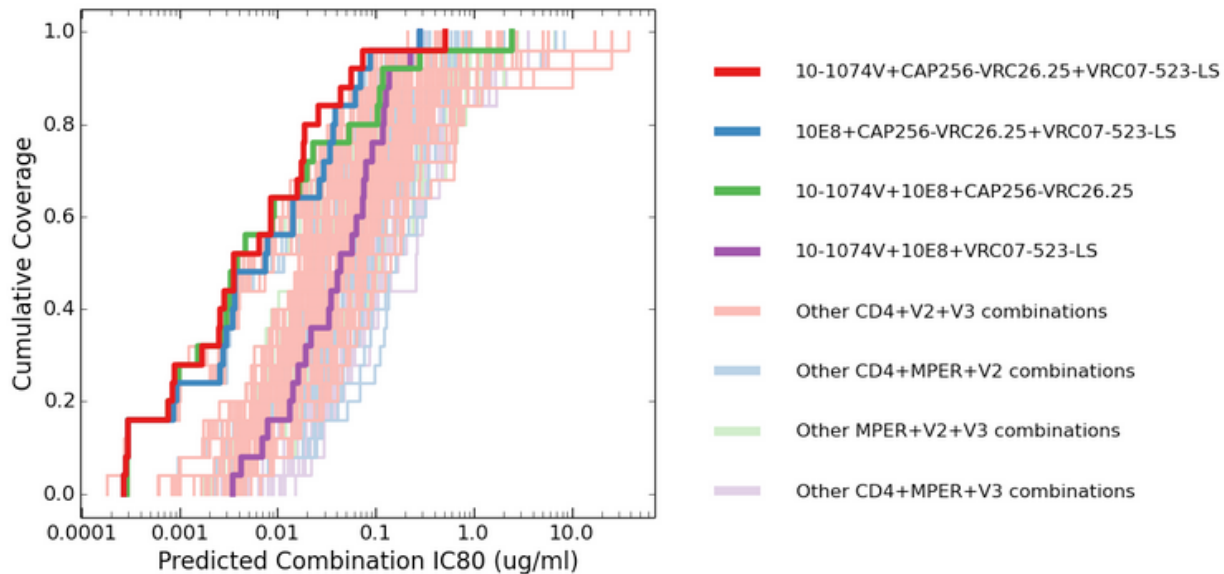
CombiNAber

Overall breadth potency ?



Single mAbs

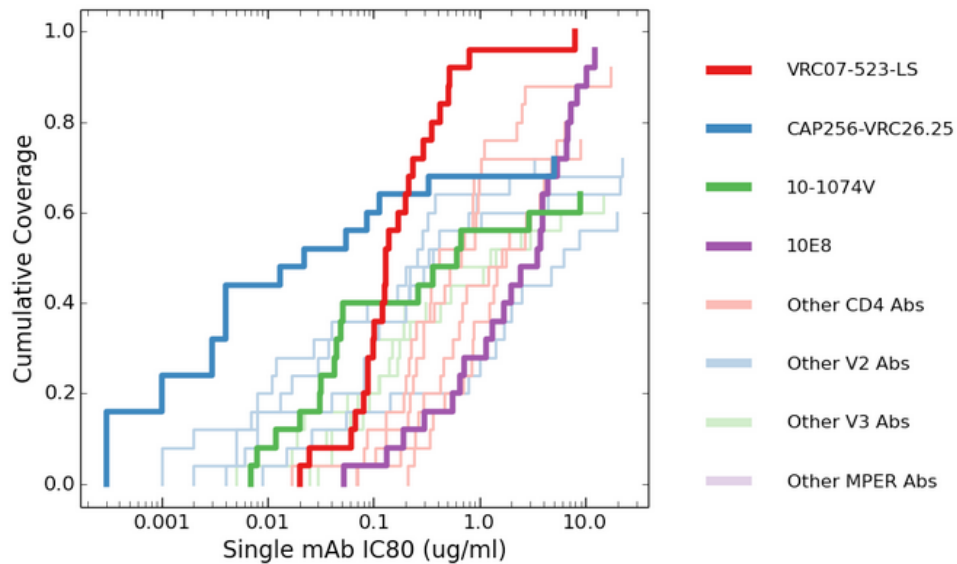
Overall breadth potency ?



3-mAb combinations

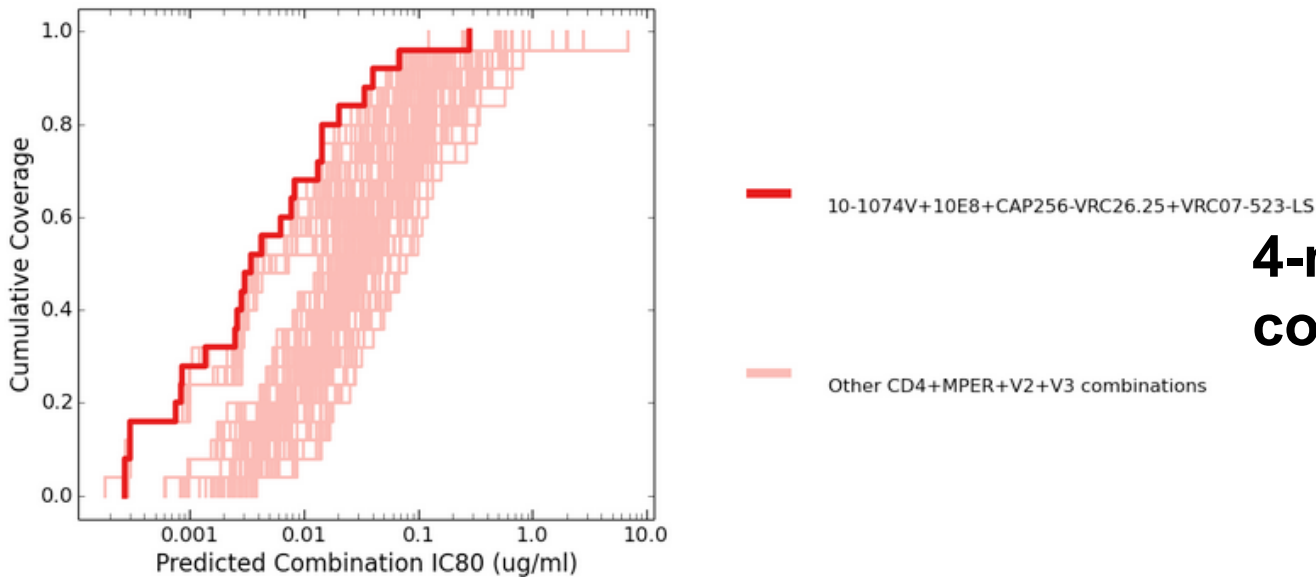
CombiNAber

Overall breadth potency ?



Single mAbs

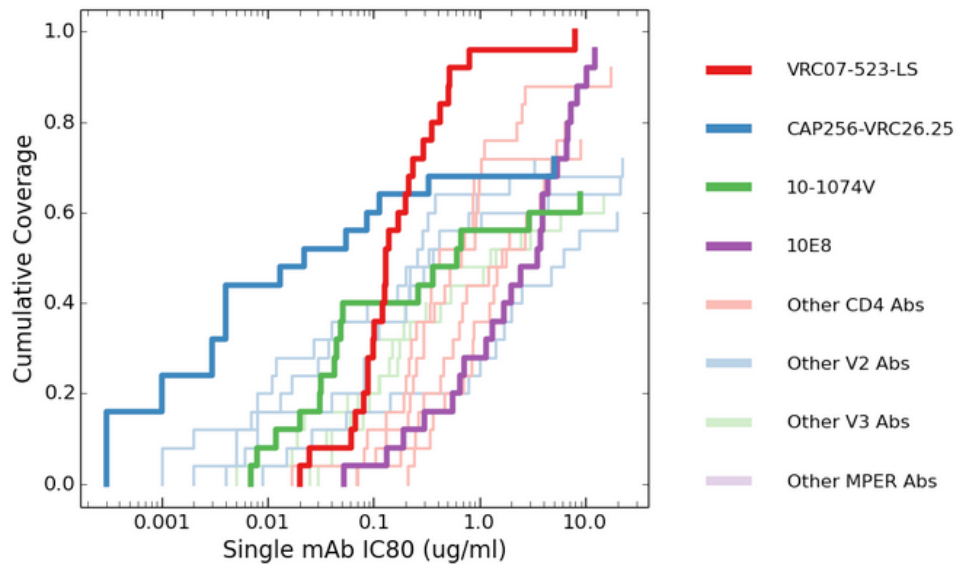
Overall breadth potency ?



4-mAb
combinations

CombiNAber

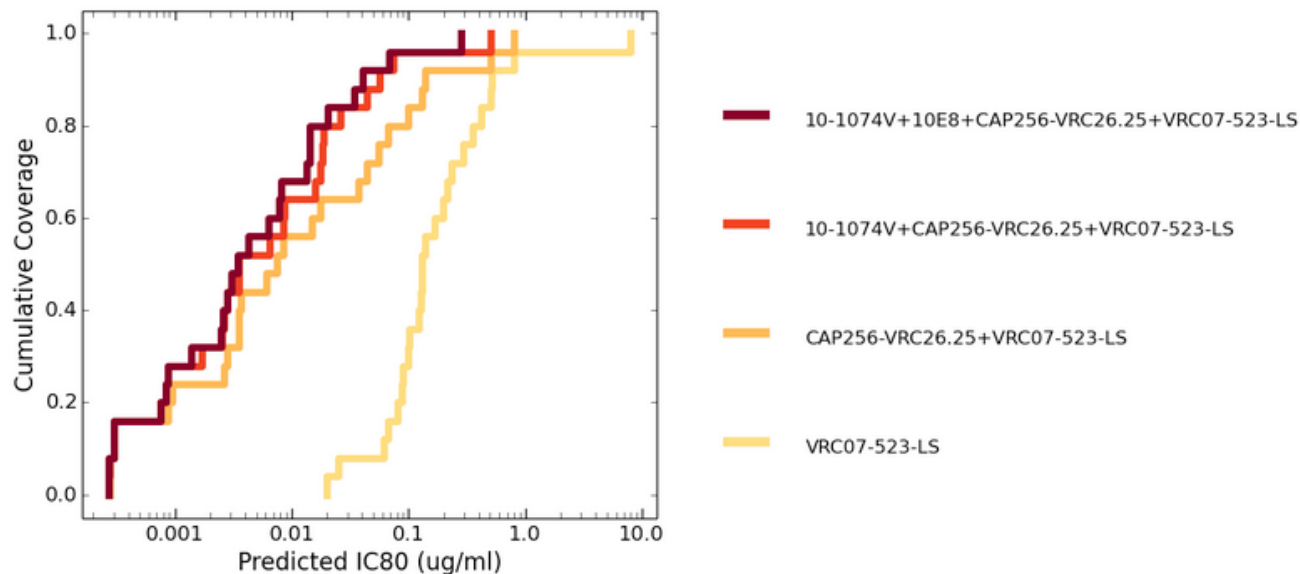
Overall breadth potency ?



Single mAbs

Wagh et al. Optimal Combinations of Broadly Neutralizing Antibodies for Prevention and Treatment of HIV-1 Clade C Infection. PLoS Pathog. 2016 12:e1005520.

Overall breadth potency ?



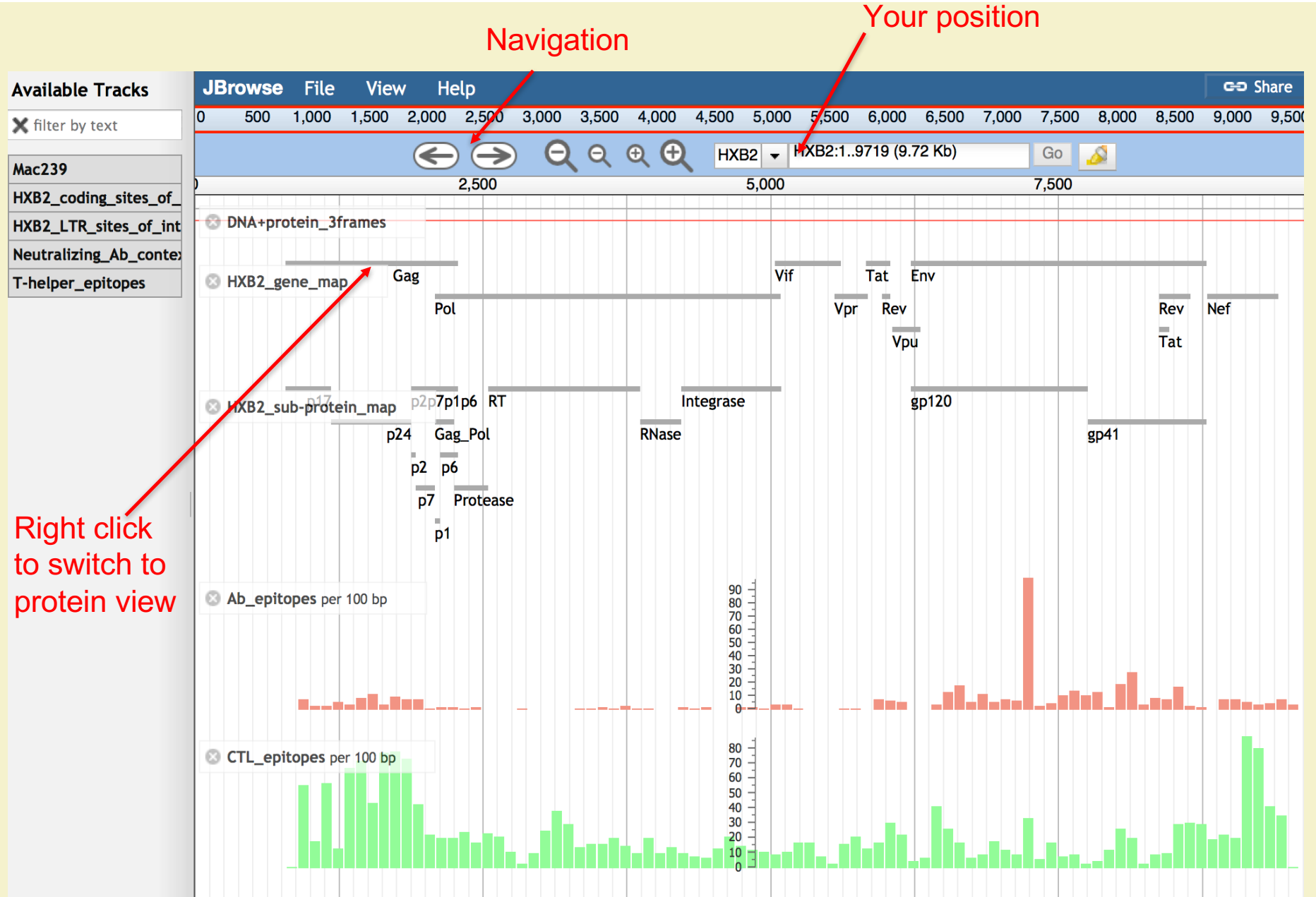
Best 4, 3, 2, 1 Combinations

HIV Genome Browser:

- A customization of [JBrowse](#) Genome Browser, built to incorporate many sources of information from our Sequence and Immunology databases.
- A one-stop source of information about HIV genome and immunological data. It retrieves the vast and diverse information available at HIV Immunology database and allow to look at the whole HIV genome and zoom in to a region of interest and see all information we have in the database about this region
 - HXB2 gene map, HXB2 sub-protein map, Mac239 map
 - Overlapping epitopes, antibody binding sites
 - HXB2 coding sites of interest (e.g. functional domains, drug resistance sites, motifs, glycosylation sites, etc.)
 - HXB2 LTR sites of interest (RNA structural elements, primer binding sites, etc.)
 - Neutralizing Ab contact residues, signatures and other NAb-associated features
 - HIV sequence variability (Entropy: M group, B clade, C clade)
 - Links to the database annotation, alignments, tools, PubMed etc.
- DNA- and Protein-level views are available

- Dreamt of by [Christian Brander](#);
- Implemented by [Shihai Feng](#);
- Help from [Jennifer Macke](#), [Brian Foley](#), [Jim Szinger](#), [Karina Yusim](#)

HIV Genome Browser: Nucleotide view



Available Tracks

X filter by text

Sub-protein_map

T-helper_epitopes

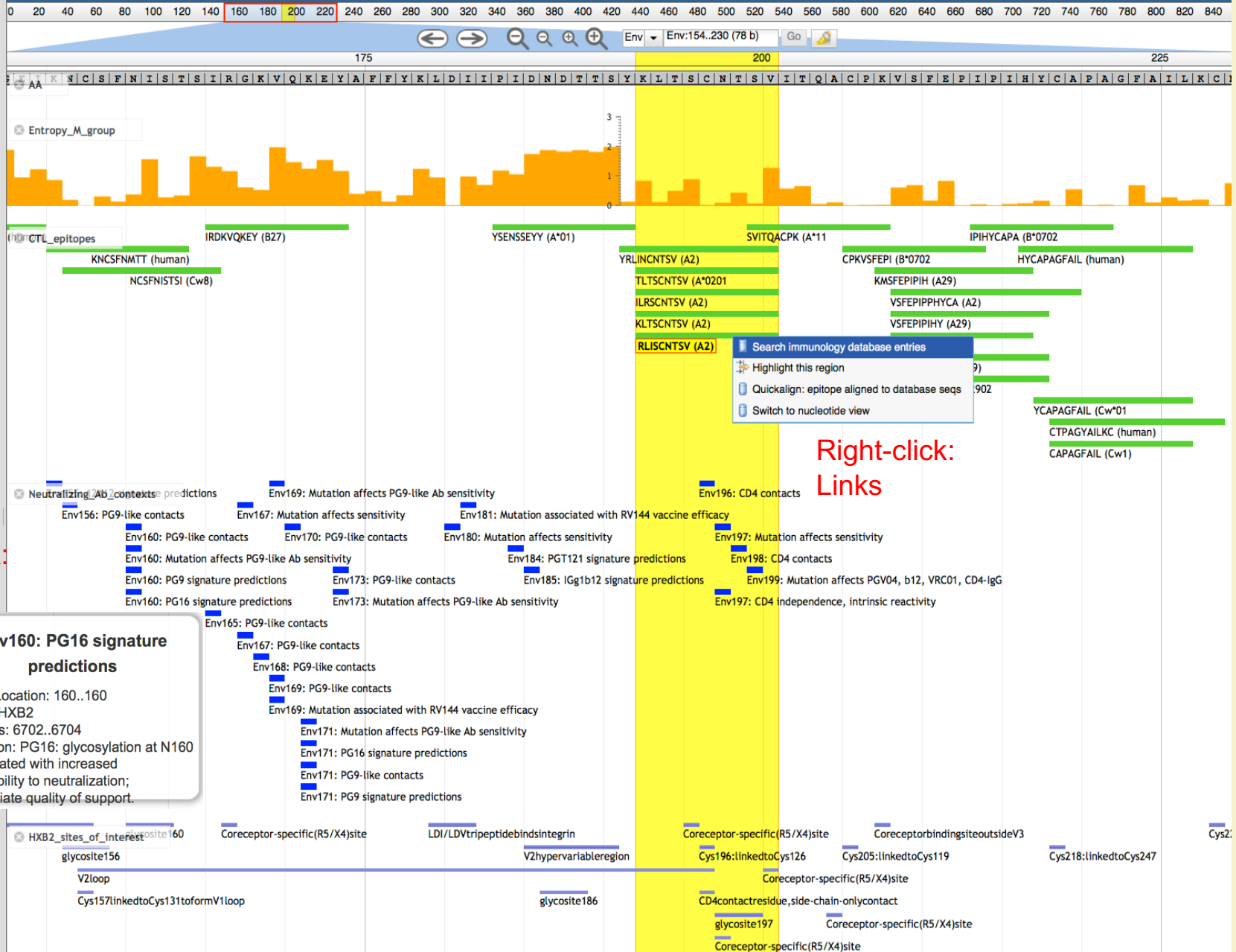
Ab_epitopes

Entropy_C_clade

Entropy_B_clade

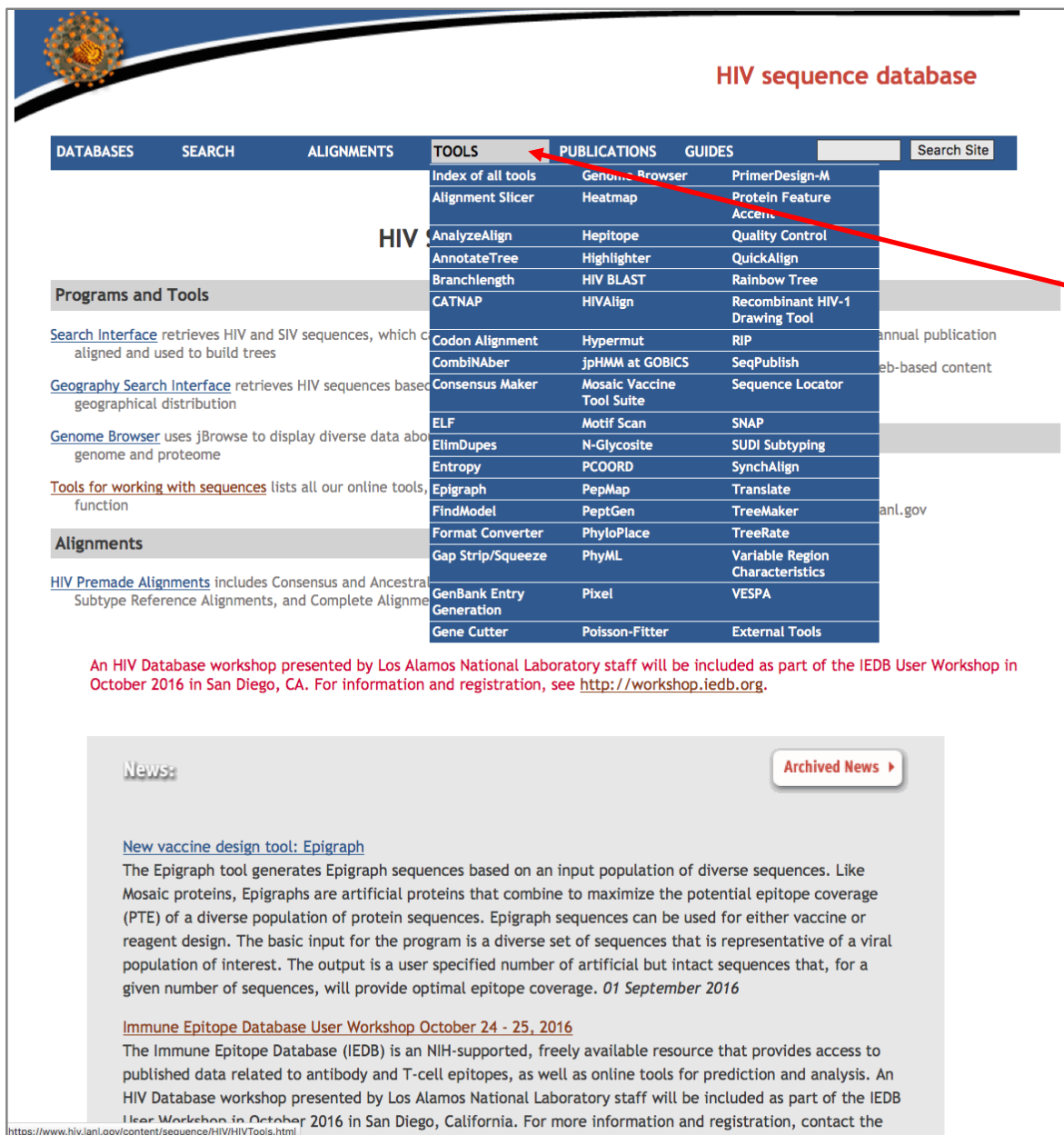
JBrowse File View Help

Share

Left-click:
DetailsRight-click:
LinksEnv160: PG16 signature
predictions

Protein Location: 160..160
Protein: HXB2
DNA_pos: 6702..6704
Annotation: PG16: glycosylation at N160
is associated with increased
susceptibility to neutralization;
intermediate quality of support.

The HIV database sequence analysis tool set



HIV sequence database

TOOLS

Index of all tools	Genome Browser	PrimerDesign-M
Alignment Slicer	Heatmap	Protein Feature
AnalyzeAlign	Hepitope	Quality Control
AnnotateTree	Highlighter	QuickAlign
Branchlength	HIV BLAST	Rainbow Tree
CATNAP	HIVAlign	Recombinant HIV-1
Codon Alignment	Hypermut	RIP
CombiNAber	jpHMM at GOBICS	SeqPublish
Consensus Maker	Mosaic Vaccine	Sequence Locator
ELF	Motif Scan	SNAP
ElimDupes	N-Glycosite	SUDI Subtyping
Entropy	PCOORD	SynchAlign
Epigraph	PepMap	Translate
FindModel	PeptGen	TreeMaker
Format Converter	PhyloPlace	TreeRate
Gap Strip/Squeeze	PhyML	Variable Region
GenBank Entry	Pixel	VESPA
Gene Cutter	Poisson-Filter	External Tools

Programs and Tools

[Search Interface](#) retrieves HIV and SIV sequences, which are aligned and used to build trees

[Geography Search Interface](#) retrieves HIV sequences based on geographical distribution

[Genome Browser](#) uses jBrowse to display diverse data about the HIV genome and proteome

[Tools for working with sequences](#) lists all our online tools by function

Alignments

[HIV Premade Alignments](#) includes Consensus and Ancestral Subtype Reference Alignments, and Complete Alignments

An HIV Database workshop presented by Los Alamos National Laboratory staff will be included as part of the IEDB User Workshop in October 2016 in San Diego, CA. For information and registration, see <http://workshop.iedb.org>.

News:

[New vaccine design tool: Epigraph](#)

The Epigraph tool generates Epigraph sequences based on an input population of diverse sequences. Like Mosaic proteins, Epigraphs are artificial proteins that combine to maximize the potential epitope coverage (PTE) of a diverse population of protein sequences. Epigraph sequences can be used for either vaccine or reagent design. The basic input for the program is a diverse set of sequences that is representative of a viral population of interest. The output is a user specified number of artificial but intact sequences that, for a given number of sequences, will provide optimal epitope coverage. 01 September 2016

[Immune Epitope Database User Workshop October 24 - 25, 2016](#)

The Immune Epitope Database (IEDB) is an NIH-supported, freely available resource that provides access to published data related to antibody and T-cell epitopes, as well as online tools for prediction and analysis. An HIV Database workshop presented by Los Alamos National Laboratory staff will be included as part of the IEDB User Workshop in October 2016 in San Diego, California. For more information and registration, contact the

<https://www.hiv.lanl.gov/content/sequence/HIV/HIVTools.html>

All tools can be accessed from the HIV sequence database

Click top level to link to full page of tools, where all >60 computational analysis tools are organized in groups by function/purpose.

Most tools have explanation pages, and sample data sets.

Many tools were inspired by user comments — please ask for more!

HIV Immunology Tools are a subset of the HIV Sequence Tools

www.hiv.lanl.gov/content/immunology/tools-links.html

HIV molecular
immunology database

Databases	Search	Tools	Products	Publications	Search Site
<div>HIV Molecular Immunology Database: Tools & Links</div>					
<div>Tools Produced by the Los Alamos National Laboratory</div> <ul style="list-style-type: none"> • CATNAP: Compile, Analyze, and Translate - Translate nucleotide sequences into amino acid sequences • HIV Genome Browser - Display HIV genome sequences • QuickAlign - Align amino acid sequences • Analyze Align - Show weblog of alignments • Alignment Slicer - Cut vertical slices from alignments • PeptGen - Generate overlap peptides • PepMap - Generate peptide maps • Motif Scan - Scan alignments for motifs <ul style="list-style-type: none"> ◦ HLA genotype/serotype ◦ HLA genotype/motif ◦ HLA supertype dictionary • Hepitope - Search for hoped-for epitopes • HLA Frequency Analysis Tool - Analyze frequencies or HLA linkage disequilibrium in a population • ELF - Epitope location finder • Sequence Locator Tool - Find the location of any HIV/SIV sequence • SeqPublish - Produce pretty alignments for publication • Heatmap - Display a table of numbers using colors to represent the numerical values • Epigraph Vaccine Suite - Design and assess Epigraphs for vaccine design • Mosaic Vaccine Suite - Design and assess polyvalent protein sequences for T-cell vaccines • N-Glycosite - Find N-linked glycosylation sites • Highlighter - Highlight matches and mismatches in a set of aligned sequences • Protein Feature Accent - View 3D graphics of HIV proteins • Variable Region Characteristics - analyzes Env variable loops and reports length, glycosolations, and net charge • All Tools - List of all software and tools in both the HIV sequence and immunology databases 					

Tools especially useful from immunologists can be accessed from the HIV Immunology “Tools” page

External Tools for Epitope Prediction

- [BIMAS HLA Peptide Binding Predictions](#) - Ranks potential n-mer peptides based on a predicted half-time of dissociation to HLA class I molecules



HIV/SIV Sequence Locator Tool

- Calculates DNA or protein fragment location relative to a reference strain
 - Available for HIV-1, SIV, HCV, and similar tools exist in HFV database
 - Such numbers, often included in the literature, are frequently incorrect

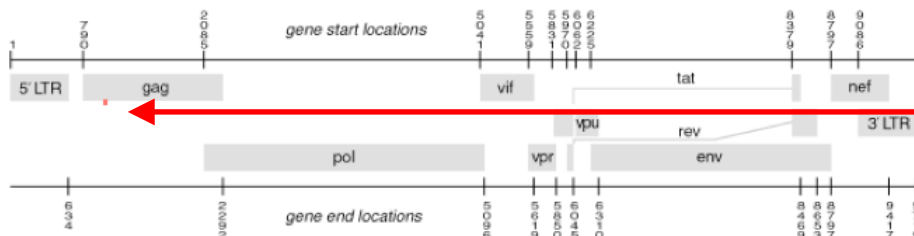
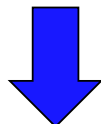
Find the location of a sequence

Sequence type ☒ Let program decide ☐ HIV ☐ SIV

Paste your input here
[Sample Input]

SLYNTVATL

Paste or type a DNA or protein sequence here.



Location in genome mapped in red.

Table of protein regions touched by query sequence. AA = amino acid, NA = nucleic acid.					
CDS	AA position relative to protein start in HXB2	AA position relative to query sequence start	AA position relative to polypeptide start in HXB2	NA position relative to CDS start in HXB2	NA position relative to HXB2 genome start
Gag	77 → 85	1 → 9	NA	229 → 255	1018 → 1044
p17	77 → 85	1 → 9	NA	229 → 255	1018 → 1044

Alignment of the query sequence to HXB2 (Similarity 100.0%):

```
Query SLYNTVATL 9
      ::::::::::
HXB2 SLYNTVATL
```

<http://www.hiv.lanl.gov/content/sequence/LOCATE/locate.html>

HIV/SIV Sequence Locator Tool

- Can also retrieve reference sequences
 - by coordinates (range of base or amino-acid positions)
 - by single position (retrieves flanking sequences)

-- OR --

Retrieve a region by its coordinates

Enter coordinates: from to (Enter '1' and 'end' to retrieve the entire region.)

Region

Retrieve ☐ Nucleotide or ☒ protein output

☐ include surrounding region

Submit

Reset

Include surrounding region

Reference Strain	Type	Region	Start	End
HXB2	pro	complete	77	85
Retrieved Sequence: SLYNTVATL				

Reference Strain	Type	Region	Start	End
HXB2	pro	complete	56	106
Retrieved Sequence: GCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIEIKDTKEALDKIEE				

50 aa long stretch

<http://www.hiv.lanl.gov/content/sequence/LOCATE/locate.html>

PepMap

www.hiv.lanl.gov/content/sequence/PepMap/pepmap.html

- Maps an input set of peptides on the query sequence
- Can be used to map epitopes, functional domains, or any protein region of interest
- Peptide name can contain any kind of useful information

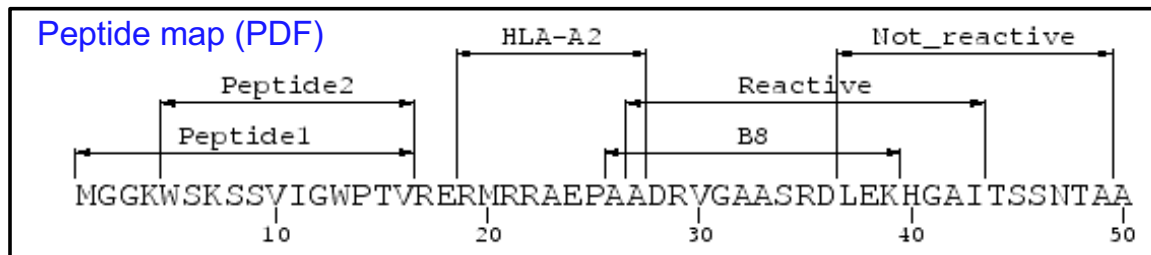
Input:

Peptide1 MGGKWSASSVIGGPTV
 Peptide2 WSKSSVIGWVTV
 HLA-A2 RMRRAEPAV
 B8 AADRGAASRDLEK
 Reactive ADRGAASRDLEKHGAI
 Not_reactive LEKHGAITSSNTA

```
>B.FR.83.HXB2_LAI_IIIB_BRU_K03455 Peptide map (FASTA)
MGGKWSKSSVIGWPTVRERMRAEPAADRGAASRDLEKHGAITSSNTAA

>Peptide1
MGGKWSASSVIGGPTV-----
>Peptide2
----WSKSSVIGWVTV-----
>HLA-A2
-----RMRRAEPAV-----
>B8
-----AADRGAASRDLEK-----
>Reactive
-----ADRGAASRDLEKHGAI-----
>Not_reactive
-----LEKHGAITSSNTA-----
```

Peptide map (PDF)



Location table

Epitope Name	Query Peptide	Reference Peptide	Protein	AA position In Protein	Polyprotein	AA position In Polyprotein	Similarity%
Peptide1	MGGKWSASSVIGGPTV	MGGKWSKSSVIGWPTV	Nef	1-16	-	-	87.5
Peptide2	WSKSSVIGWVTV	WSKSSVIGWPTV	Nef	5-16	-	-	91.7
HLA-A2	RMRRAEPAV	RMRRAEPA	Nef	19-27	-	-	88.9
B8	AADRGAASRDLEK	AADRGAASRDLEK	Nef	26-39	-	-	100.0
Reactive	ADRGAASRDLEKHGAI	ADRGAASRDLEKHGAI	Nef	27-43	-	-	100.0
Not_reactive	LEKHGAITSSNTA	LEKHGAITSSNTA	Nef	37-49	-	-	100.0

PeptGen

- Generates overlapping peptides for any protein sequence
- Takes alignment as an input and removes duplicate peptides

```
Seq1      HIVWASRELERFAVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIEVKDTKEALEKIEEEQNKSK
Seq2      HLVWASRELERFALNPGLLETSEGCKQIIKQLQPALQTGTEELRSLYNTVATLYCVHEKIEVRDTKEALDKIEEEQNKSQ
Seq3      HLVWASRELERFALNPDLLTAEGCQQIMGQLQPALQTGTEELRSLFNTVATLYCVHQRIEVKDTKEALEEVEKIQKKSQ
```

```
HIVWASRELERFAVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIEVKDTKEALEKIEEEQNKSK
HIVWASRELERFAVNPGL CON_B (18)
-L-----L---- CON_C
-L-----L--D- CON_G

  LERFAVNPGLLETSEGCR CON_B (18)
  -----L-----K CON_C
  -----L--D---A---Q CON_G

    GLLETSEGCRQILGQLQP CON_B (18)
    -----K--IK---- CON_C
    D-----A---Q--M---- CON_G

      CRQILGQLQPSLQTGSEE CON_B (18)
      -K--IK----A----T-- CON_C
      -Q--M-----A----T-- CON_G

        QPSLQTGSEELRSLYNTV CON_B (18)
        --A----T----- CON_C
        --A----T-----F--- CON_G

          EELRSLYNTVATLYCVHQ CON_B (18)
          -----E CON_C
          -----F----- CON_G

            TVATLYCVHQRIEVKDTK CON_B (18)
            -----EK---R--- CON_C
            ----- CON_G

              HQRIEVKDTKEALEKIEE CON_B (18)
              -EK---R-----D--- CON_C
              -----EV-K CON_G

                TKEALEKIEEEQNKSK CON_B (16)
                -----D-----Q CON_C
                -----EV-KI-K--Q CON_G
```

```
1 HIVWASRELERFAVNPGL 1 s1 1 s1 - -
2 HLVWASRELERFALNPGL 1 s2 1 - s2 -
3 HLVWASRELERFALNPDL 1 s3 1 - - s3

4 LERFAVNPGLLETSEGCR 2 s1 1 s1 - -
5 LERFALNPGLLETSEGCK 2 s2 1 - s2 -
6 LERFALNPDLLTAEGCQ 2 s3 1 - - s3

7 GLLETSEGCRQILGQLQP 3 s1 1 s1 - -
8 GLLETSEGCKQIIKQLQP 3 s2 1 - s2 -
9 DLLETAGCQQIMGQLQP 3 s3 1 - - s3

10 CRQILGQLQPSLQTGSEE 4 s1 1 s1 - -
11 CKQIIKQLQPALQTGTEE 4 s2 1 - s2 -
12 CQQIMGQLQPALQTGTEE 4 s3 1 - - s3

13 QPSLQTGSEELRSLYNTV 5 s1 1 s1 - -
14 QPALQTGTEELRSLYNTV 5 s2 1 - s2 -
15 QPALQTGTEELRSLFNTV 5 s3 1 - - s3

16 EELRSLYNTVATLYCVHQ 6 s1 1 s1 - -
17 EELRSLYNTVATLYCVHE 6 s2 1 - s2 -
18 EELRSLFNTVATLYCVHQ 6 s3 1 - - s3

19 TVATLYCVHQRIEVKDTK 7 s1&s3 2 s1 - s3
20 TVATLYCVHEKIEVRDTK 7 s2 1 - s2 -

21 HQRIEVKDTKEALEKIEE 8 s1 1 s1 - -
22 HEKIEVRDTKEALDKIEE 8 s2 1 - s2 -
23 HQRIEVKDTKEALEEVEK 8 s3 1 - - s3
```

QuickAlign

- Aligns query sequence to an alignment, creates WebLogos, calculates frequency by position, tallies variants in an alignment
- Can be used to align epitopes, functional domains, or any protein or any region of interest
- Shows results by groupings (subtypes for example) and all groups together

Query:	SLYNTVATL
Query Length:	9
HXB2 Location:	Gag 77-85 = p17 77-85
Alignment:	GAG, 458 sequences
Summarize	
Query	SLYNTVATL
A1.KE.86.ML170	--F-----
A1.KE.94.Q23	--F-----
A1.SE.94.SE7253	--F----V-
A1.SE.94.SE7535	-----
A1.SE.95.SE8538	-----
A1.SE.95.SE8891	-----
A1.SE.95.UGSE8131	-----
A1.TZ.97.97TZ03	--F----V-

Summary for subtype A

Variant	Count	Percent
SLYNTVATL		
--F-----	11	47.83
-----	7	30.43
--F--I-V-	1	4.35
--F----V-	1	4.35
-----V-	1	4.35
----L----	1	4.35
--F-A--V-	1	4.35

Variant
frequency
summary

Total sequences = 23

Number of variants = 7

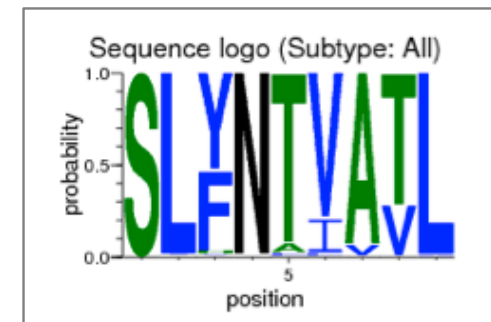
Frequency by position

[See full raw counts](#)

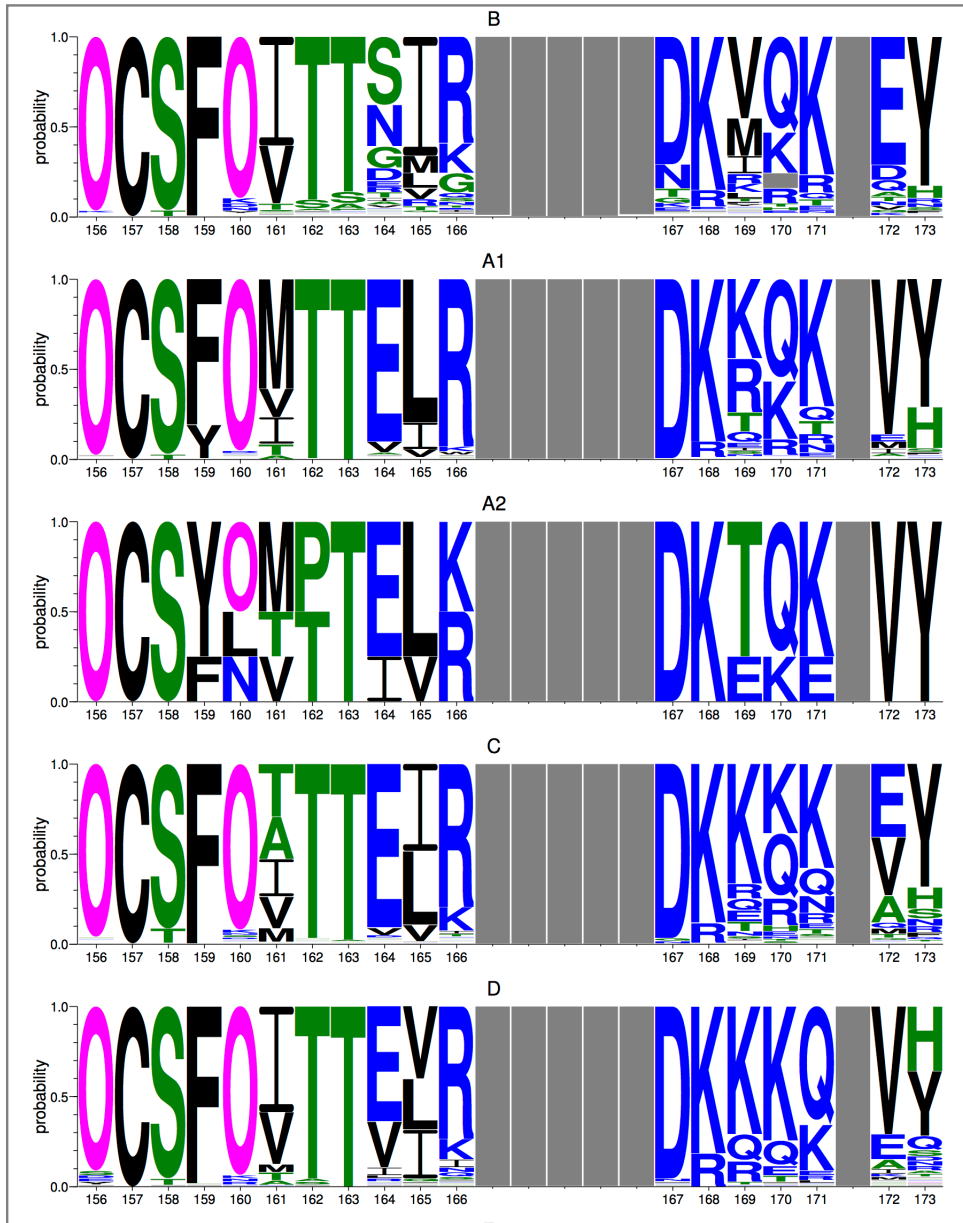
[Go to top](#)

cutoff: 95%

Position	Percentage and raw count of non-gap			Non-gap/total (percentage)
1	S: 99.90% (3113)	other: 0.10% (3)		3116/3119 (100.00%)
2	L: 98.90% (3068)	other: 1.10% (34)		3102/3119 (99.55%)
3	Y: 52.71% (1633)	F: 43.77% (1356)	other: 3.52% (109)	3098/3119 (99.42%)
4	N: 99.68% (3104)	other: 0.32% (10)		3114/3119 (99.94%)
5	T: 92.86% (2887)	A: 5.05% (157)	other: 2.09% (65)	3109/3119 (99.78%)
6	V: 79.35% (2448)	I: 18.15% (560)	other: 2.50% (77)	3085/3119 (99.01%)
7	A: 92.95% (2889)	V: 6.53% (203)	other: 0.51% (16)	3108/3119 (99.74%)
8	T: 72.52% (2254)	V: 27.06% (841)	other: 0.42% (13)	3108/3119 (99.74%)
9	L: 99.00% (3078)	other: 1.00% (31)		3109/3119 (99.78%)



MAb PG9 binding regions, Env 156-173,
bNAb PG9 contact region



AnalyzeAlign

- New tool similar to QuickAlign, but takes sequence positions/range (including discontinuous) to analyze in an alignment
- Has many analysing options:
 - WebLogo specifications
 - Frequency cutoffs
 - Choice of the master sequence to find variants
 - User-specified color scheme
 - Combining multiple logos on a page
 - Showing potential N-linked glycosylation sites (Nx[ST], denoted as **0**)

AnalyzeAlign



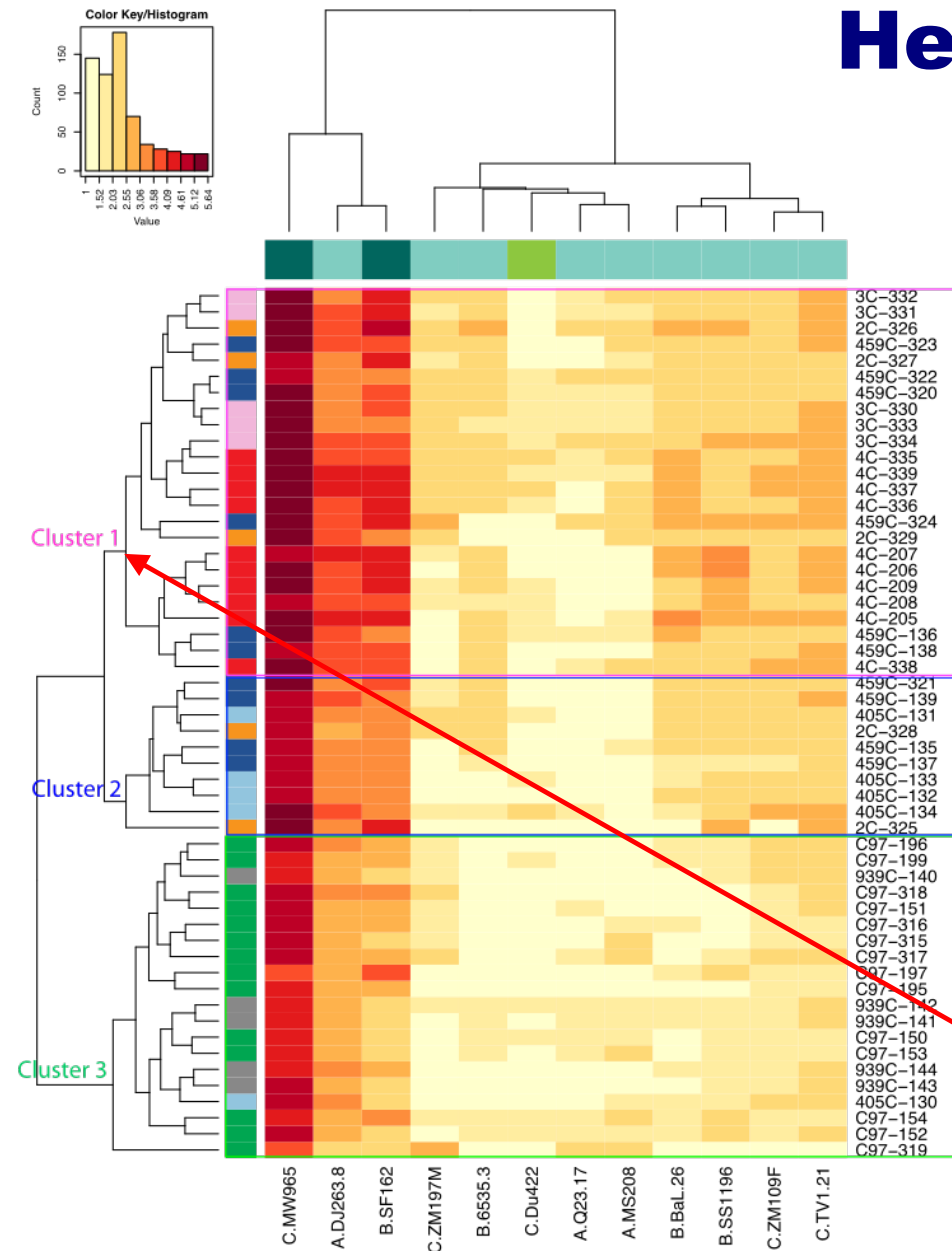
Transmitted HIV virus

- ☐ Longitudinal samples over time
- ☐ Discontinuous positions under apparent immune selection
- ☐ Only differences from transmitted virus are shown
- ☐ Colors indicate amino acid charge categories

Virus 3 years post-infection

Figure from Hraber et al. *Viruses* 2015

Heatmap



- Two-dimensional clustering analysis
- A graphical way of displaying a table of numbers by using colors to represent numerical values.
- Strategy borrowed from the gene expression array literature to organize and visualize neutralization data, but is also useful for other complex data
- *Example: (Bricault et al, 2015, J Virol)*
 - Rows: ID50s in guinea pigs vaccinated by 4 different strains and combinations of 2, 3, and 4 strains (2C, 3C, 4C)
 - Columns: Tier 1A, 1B test Envs
 - Higher intensity color – higher ID50s
 - Vertical bar – animals colored by vaccine
 - Horizontal bar – Envs colored by neutralization tier
 - Animals vaccinated by 4 strains in combination cluster together on the top and have highest ID50s

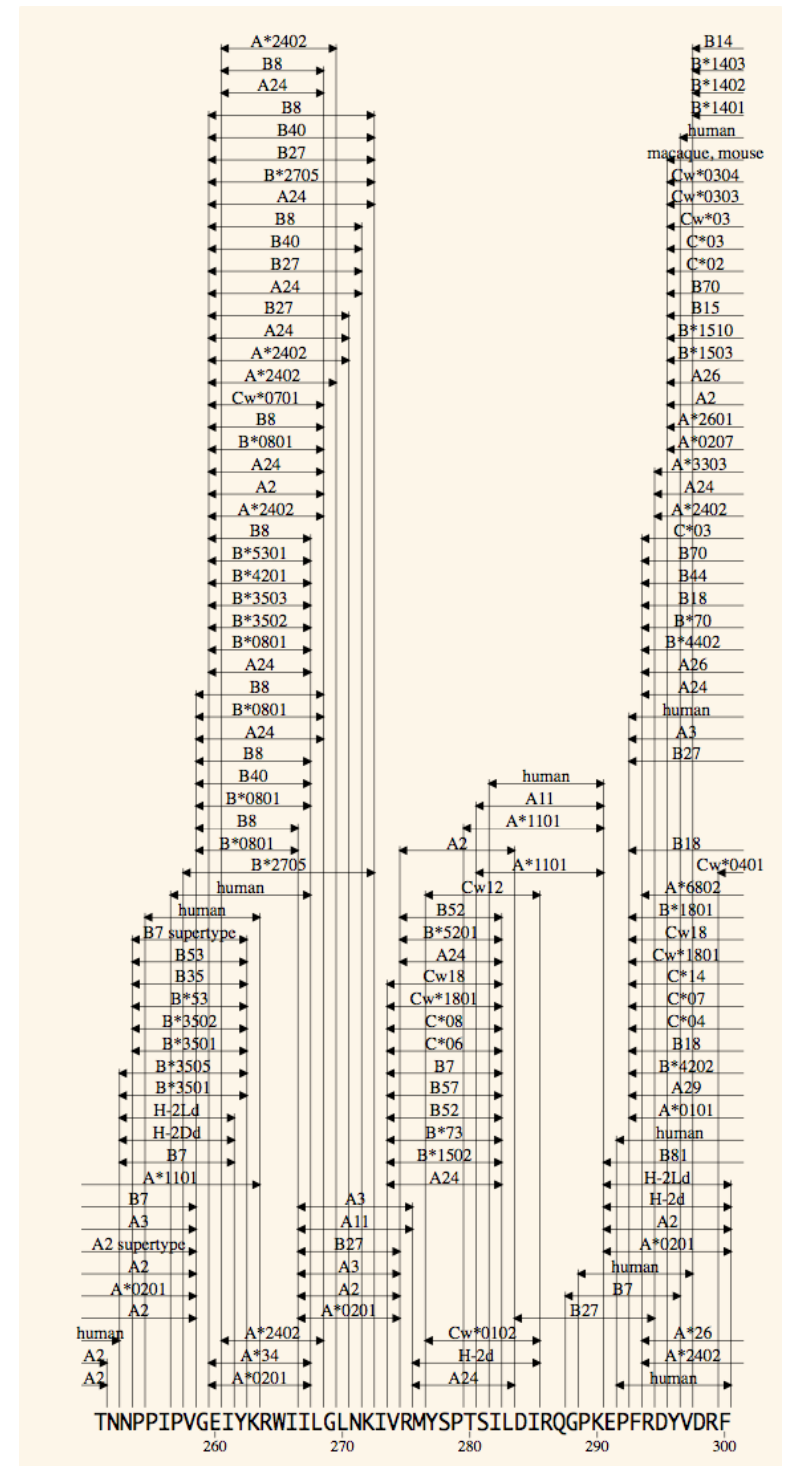
Column Colors
 Tier_1A Tier_1B Tier_2

Row Colors
 2C 3C 405C 459C 4C 939C C97

p17 CTL/CD8+ Epitope Map

- Epitopes up to 14 aa long are mapped on HXB2
- HXB2 sequence may differ
- Epitopes with identical boundaries and HLA fields are included in the maps only once
- The epitope maps are interactive!
 - Clicking on an epitope leads to the epitope entry

www.hiv.lanl.gov/content/immunology/maps/maps.html



CTL/CD8+ Epitope Summary (B-list)

- A comprehensive list of all unique epitopes in the database (including with unknown HLA, boundaries not fully defined...)
- Similar lists for Helper epitopes and linear Ab binding sites
- Unlike epitope maps that show epitope locations, each epitope sequence is shown

Epitope	Protein	HXB2 Location	Subtype	Species	HLA
MGARASVLSG	p17	1-10	CRF01_AE	human	
ASVLSGGEL	p17	5-13	B	human	
ASILRGGKLDK	p17	5-15	C	human	
SVLSGGQLDR	p17	6-15	B	human	A11
LSGGELDRWEK	p17	8-18		macaque	
GELDRWEKI	p17	11-19	B	human	B*4002, B40
GQLDRWEKI	p17	11-19	B	human	
GKLDSEKIRLR	p17	11-22	A, CRF01_AE, CRF02_AG	human	

www.hiv.lanl.gov/content/immunology/tables/ctl_summary.html

Best-defined CTL/CD8+ Epitope Summary (A-list)

- Experimentally validated optimal epitopes with known HLA presenting molecules
- Defined/curated by Christian Brander and colleagues

Epitope	Protein	HXB2 Location	Subtype	Species	HLA
GELDRWEKI	p17	11-19		human	B*4002
KIRLRPGGK	p17	18-26		human	A*0301
IRLRPGGKK	p17	19-27	B	human	B*2705
RLRPGGKKK	p17	20-28		human	A*0301
RLRPGGKKKY	p17	20-29	B	human	A*0301
GGKKKYKLK	p17	24-32	B	human	B*0801
KYKLKHIVW	p17	28-36	B	human	A*2402
HLVWASREL	p17	33-41		human	Cw*0804

www.hiv.lanl.gov/content/immunology/tables/optimal_ctl_summary.html

Epitope variants and escape mutations

- Experimental epitope variants from the literature
 - Search interfaces
 - Summary tables (~3500 CTL epitope variants)
- HLA associated HIV polymorphisms (Zabrina Brumme, Bruce Walker)
 - Database review and a table

www.hiv.lanl.gov/content/immunology/pdf/2010/escape_article_supplement.html

CTL/CD8+ Search (www.hiv.lanl.gov/content/immunology/ctl_search)

HIV protein	Proteins with defined epitopes - ALL - p17 p17-p24 p24 p24-p2p7p1p6	Proteins with undefined epitopes - ALL - Gag Gag/Pol Pol Vif
HXB2 location	<input type="text"/> - <input type="text"/>	Results overlap with query location
Epitope	ISPRTLNAW	Results contain query sequence
Epitope name	<input type="text"/>	
Record number	<input type="text"/>	
Subtype	- ALL -	
Immunogen	- ALL - computer prediction HIV-1 and GBV-C co-infection HIV-1 and HCV co-infection HIV-1 exposed seronegative HIV-1 infected monocyte-derived HIV-1 infection	
Vaccine details	Vaccine type Vaccine strain if Immunogen is Vaccine Vaccine component Adjuvant	- ALL - - ALL - - ALL - - ALL -
Species	- ALL -	
MHC/HLA	- ALL - A*01 A*0101 A*02 A*0201 A*02.01 A*020101	
Author	Pillay	<input checked="" type="checkbox"/> First <input type="checkbox"/> Last
Country	- ALL -	
Keywords	- ALL - acute/early infection adjuvant comparison antagonism antibody binding site definition and exposure assay development, comparison, standardization, improvement autologous responses	
Note	<input type="text"/>	

- Search by HIV protein, Epitope Sequence, Subtype, Immunogen, Vaccine Details, Species, presenting MHC/HLA, Author, Country, Keywords

- Search on epitope location and find fuzzy matches, overlaps and embedded epitopes

- Search examples:

□ *Example:*

- SLYNTVATL – 285 entries
- Narrow the search with keyword “escape” – 35 entries

**Search for ISPRTLNAW
With the first author Pillay**

Search

Reset

Click for [Search Help](#)

Search CTL/CD8+ T-Cell Epitope Database

Found 1 matching record:

Displaying record number 53832

HXB2 Location	p24(15-23)
Author Location	Gag(147-155)
Epitope	ISPRTLNAW
Subtype	C
Species (MHC/HLA)	human(B57)
Immunogen	HIV-1 infection
Donor MHC/HLA	A*3001, A*66, B*4201, B*5802, Cw*0602, Cw*1701; A*66, A*68, B*57, B*5802, Cw*0602, Cw*0701
Country	South Africa
Experimental methods	CD8 T-cell Elispot - IFN γ
Keywords	epitope processing, responses in children, mother-to-infant transmission, escape, acute/early infection

[Link to Epitope Maps](#)

[Link to Epitope Alignment](#)

[Variant details with annotator's notes](#)

[p24 Epitope Map](#)

[Epitope Alignment](#)

[Show epitope variants](#)

Additional information provided in the entry:

- Location, Donor MHC/HLA, experimental methods, Notes
- Link to all entries for a reference
- PubMed links to papers
- Link to Epitope Maps
- Link to Epitope Alignment (aligned to large set of seq.)
- Epitope variants if studied in the paper

Notes

- HIV-specific CTLs in infants were shown to be able to select for viral escape variants early in life, despite a lack of escape with the same CTL specificity in the mother. Infant CTL responses may be compromised by transmission of escape variants that arose in the mother and also those that arose in the father, if the father was the source of the mother's infection.
- ISPRTLNAW is the C consensus form of the epitope and was the autologous form in the mother, and was transmitted to her infant. By 33 weeks a new dominant form of the epitope had emerged in the infant, mSPRTLNAW, and two additional variants had arisen, one with a substitution proximal to the epitope, pISPRTLNAW, and ISPRTLNAW.

References

Pillay2005 Thillagavathie Pillay, Hua-Tang Zhang, Jan W. Drijfhout, Nicola Robinson, Helen Brown, Munira Khan, Jagadesa Moodley, Miriam Adhikari, Katja Pfafferott, Margaret E. Feeney, Anne St. John, Edward C. Holmes, Hoosen M. Coovadia, Paul Klenerman, Philip J. R. Goulder, and Rodney E. Phillips. Unique Acquisition of Cytotoxic T-Lymphocyte Escape Mutants in Infant Human Immunodeficiency Virus Type 1 Infection. *J. Virol.*, 79(18):12100-12105, Sep 2005. PubMed ID: [16140787](#). [Show all entries for this paper.](#)

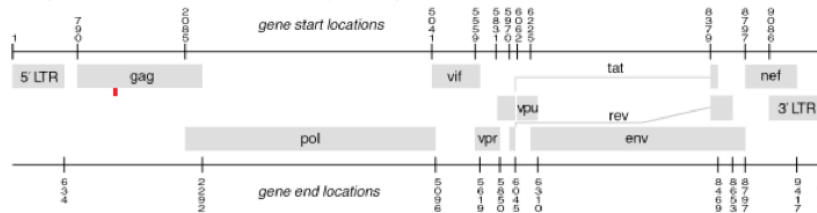
Epitope Alignment

Also available as a separate tool QuickAlign

www.hiv.lanl.gov/content/sequence/QUICK_ALIGNv2/QuickAlign.html

Genome map:

Query location(s) shown as colored bar(s) in map.



Summary & analysis:

Query: epitope

Query sequence	ISPRTLNAW
Query length	9
HXB2 Location	● genome: 1228→1254, region: Gag 147→155
Alignment used	LANL HIV1 Gag Amino acid Filtered web

☒ Summarize All
 ☐ Summarize By Subtype (major subtype only ☒)
 ☐ Find Other Matches

Alignment slice:

alignment below in [format](#)

"-" = identity to query sequence

"." = gap in sequence

"Red name" = perfect identity to query sequence

```

                                epitope ISPRTLNAW
B.FR.83.HXB2_LAI_IIIB_BRU.K03455 -----
A1.CA.x.BCCFE_HOMER_HIV_GAG_3062.EU242119 L-----
A1.CD.02.02CD_KTB035.AM000055 -----
A1.CD.97.97CD_KCC2.AM000053 L-----
A1.CD.97.97CD_KTB13.AM000054 L-----
A1.CH.03.HIV_CH_BID_V3538_2003.JQ403028 L-----
A1.CH.04.pBV23.KJ689262 L-----
A1.CH.05.pBV26.KJ689264 F-----
A1.CH.08.pBV20.KJ689259 L-----
A1.CH.09.pBV32.KJ689270 M-----
A1.CH.10.pBV17.KJ689256 M-----
A1.CH.11.pBV13.KJ689253 -----
A1.CH.11.pBV22.KJ689261 L-----
A1.CH.11.pBV48.KJ689279 L-----
A1.CH.12.pBV58.KJ689285 L-----
A1.CM.06.BS02.JX244900 L-----
A1.CM.07.46_10.KP718918 L-----
A1.CM.07.BS10.JX244906 L-----
A1.CM.08.886_24.KP718928 L-----
A1.CN.00.00CNLN14.EF122512 V-----
A1.CY.04.CY009.EU673416 L-----
A1.CY.05.CY012.EU673418 L-----
A1.CY.05.CY021.FJ388892 -A-KA-EG-
A1.CY.05.CY051.FJ388903 L-----
    
```

Displaying record number 53832

Variant details

HXB2 Location	p24(15-23)	p24 Epitope Map
Epitope	ISPRTLNAW	Epitope Alignment
Variants	mSPRTLNAW	escape documented in this paper
	lSPRTLNAW	diminished response
	p lSPRTLNAW	not determined
Species (MHC/HLA)	human(B57)	

Link back to epitope entry

Variant Details

Showing all 3 variants.

Variant ID.	1413
Epitope Seq.	ISPRTLNAW
Variant Seq.	mSPRTLNAW
Mutations	I/M
Epitope Location	I1M
HXB2 Location	I15M
Mutation Type	E: escape documented in this paper
Method	CD8 T-cell Elispot - IFN γ , Sequence
Note	This is de novo variant seen in infant by week 33 of age. The index peptide was recognized, but not the variant.

Variant ID.	1414
Epitope Seq.	ISPRTLNAW
Variant Seq.	lSPRTLNAW
Mutations	I/L
Epitope Location	I1L
HXB2 Location	I15L
Mutation Type	DR: diminished response
Method	CD8 T-cell Elispot - IFN γ , Sequence

Mutation type

Note describing why the variant was designated as a particular mutation type

Mutation type examples:

- ☐ E escape
- ☐ IE inferred escape
- ☐ DR diminished response
- ☐ SF susceptible form
- ☐ etc...

ELF (Epitope Location Finder)

ELF
Epitope Location Finder

Purpose: search a submitted protein sequence for (1) known epitopes from our immunology databases, (2) epitopes predicted by consensus binding motifs, and (3) epitopes predicted by the IEDB binding algorithm. For details see [ELF Explanation](#).

Input

Paste [protein sequence](#) <50 amino acids, raw format

Options

Show [known epitopes](#) ☒ from CTL and Helper databases

Find potential epitopes ☒ based on [anchor residues](#)

Choose [HLA\(s\)](#) (Class I and Class II)
Use control-click for multiple selection

By genotype

A*3004
A*3101
A*3201
A*3303
A*6601
A*6801
A*6802

By serotype

A33(19)
A69(28)
A68(28)
A30(19)
A66(10)
A1
A2

Find potential epitopes ☒ based on [IEDB binding predictions](#)

Choose [HLA\(s\) or MHC\(s\)](#) (synchronized with genotype selections above)

HLA Class I

A*6611
A*6612
A*6613
A*6614
A*6615
A*6801
A*6802

HLA Class II

DRB3*0224
DRB3*0225
DRB3*0301
DRB3*0303
DRB4*0101
DRB4*0103
DRB5*0101

Animal MHC Class I

chimpanzee

Patr-A*0101
Patr-A*0201
Patr-A*0301
Patr-A*0302
Patr-A*0401
Patr-A*0402

Animal MHC Class II

mouse

H2-IAb
H2-IAd
H2-IEd

Display binders ☒ Show best binder(s) per MHC

☐ Show below [percentile rank](#) (1-100) per MHC

E-mail result ☐ Predictions are slow. For more than a few HLAs/MHCs, we recommend e-mailed result.

HLA selection is
synchronized between
2 analysis options

You can choose
how many top binders
to show per MHC,
or use a binding
percentile rank cutoff


■ **ELF** helps identify potential T cell epitopes in a reactive peptide from a person with known HLA type by

- Highlighting appropriate HLA anchor motifs in the peptide
- Aligning all known epitopes embedded in the peptide from the database to your query sequence, with links to epitope entries
- Finding potential epitopes based on Immune Epitope Database (IEDB) binding predictions <http://www.immuneepitope.org/>

■ We also have **MotifScan** tool that shows HLA binding and custom motifs on the sequence alignment

ELF (reported epitopes in HIV database)




Epitopes from our CTL database aligned to your query sequence

Bold **red** letters indicate residues that differ from the query sequence. The symbol  means the HLA of the epitope matches one of your submitted HLAs. Click on the epitope to see full database entry. Click on "align" to align the epitope to the sequence database via QuickAlign.

Epitopes shown here are completely within the bounds of your query. Epitopes that overlap the ends of your query are included in the "View database records" links above.

Download this alignment in format

DTVLEDMNLPGRWKPKMIG

[DTVLE**EM**NL](#) A*6802 [align](#)  Epitopes
[DTVLE**I**NL](#) A*6802 [align](#)  matching
[DTVLE**EW**NL](#) A*6802 [align](#)  requested HLAs
[DTVLE**EM**NL](#) A68 [align](#)
[DTVLE**EM**NL](#) A28 [align](#)
[DTVLEDMNLP](#) [align](#)

Clicking on an epitope takes you to respective CTL or Helper epitope Database entries

[E**EM**NLPGRW](#) B44 [align](#)
[E**E**I**N**LPGKW B44 \[align\]\(#\)
\[E**EM**NLPGRW\]\(#\) B*4402 \[align\]\(#\)
\[E**EM**NLPGRW\]\(#\) B*4403 \[align\]\(#\)
\[E**EM**NLPGRW\]\(#\) B18,B40,B44 \[align\]\(#\)
\[EDMNLPGRW\]\(#\) \[align\]\(#\)
\[E**EM**NLPGRW\]\(#\) B*44 \[align\]\(#\)
\[E**E**I**N**LPGKW B*4403 \\[align\\]\\(#\\)
\\[E**EM**NLPGRW\\]\\(#\\) \\[align\\]\\(#\\)
\\[LPGRWKPKMI\\]\\(#\\) Cw3 \\[align\\]\\(#\\)
\\[LPGRWKPKMI\\]\\(#\\) B7 \\[align\\]\\(#\\)\]\(#\)](#)

Clicking on the "align" button takes you to "QuickAlign" for that epitope

ELF (predicted MHC binding)

Potential epitopes based on anchor residues

These peptides have C-terminal anchor residues, highlighted in **blue**, and internal anchors highlighted in **magenta**. These anchor residues match one or more motifs associated with the submitted HLA.

[Download](#) this alignment in format **table**

```
DTVLEDMNLPGRWKPKMIG
DTVLEDMNL (A*0205 .....[L])
DTVLEDMNL (A*6802 .[TV].....[VL])
TVLEDMNLP (A*0206 .[VQ].....)
LEDMNLPGR (DRB5*0101,DRB5*0101 [FYLM]..[QVIM]....[RK])
```

Motifscan

Potential epitopes based on IEDB binding predictions

Top binders for each MHC are highlighted in **blue**.

Prediction method: IEDB recommended

Low percentile = good binders

Show up to 1 binder(s) per MHC

Class I

Selected allele(s): A*6802, B*1501

[Download](#) this alignment in format **table**

DTVLEDMNLPGRWKPKMIG (Click MHC to see full list of IEDB predictions for that MHC)

DMNLPGRW	B*1501 (26)
MNLPGRWK	A*6802 (3.0)

Class II

Selected allele(s): DRB5*0101

[Download](#) this alignment in format **table**

DTVLEDMNLPGRWKPKMIG (Click MHC to see full list of IEDB predictions for that MHC)

TVLEDMNLPGRWKPK	DRB5*0101 (17.17)
---------------------------------	-----------------------------------

IEDB binding predictions

Clicking on MHC links to the full list of IEDB predictions for that MHC (see next table)

Potential epitopes based on IEDB database MHC binding predictions

IEDB Analysis Resource

[Home](#)[Help](#)[Example](#)[Reference](#)[Download](#)[Contact](#)

MHC-I binding predictions - Prediction Results

Input Sequences

#	Name	Sequence
1	sequence 1	DTVLEDMNLPGRWKPKMIG

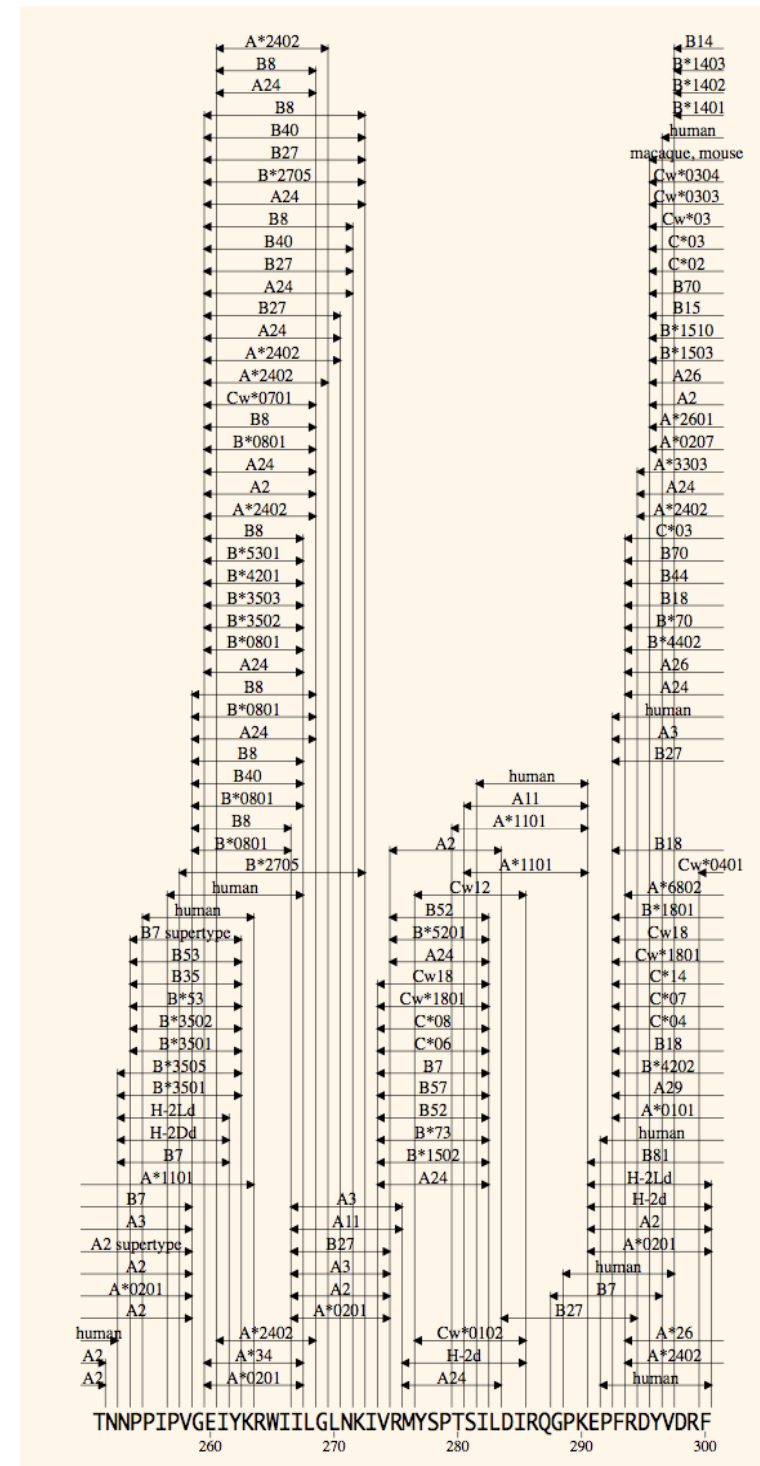
Prediction method: IEDB recommended | Low percentile = good binders

Check to expanded the result: ☐

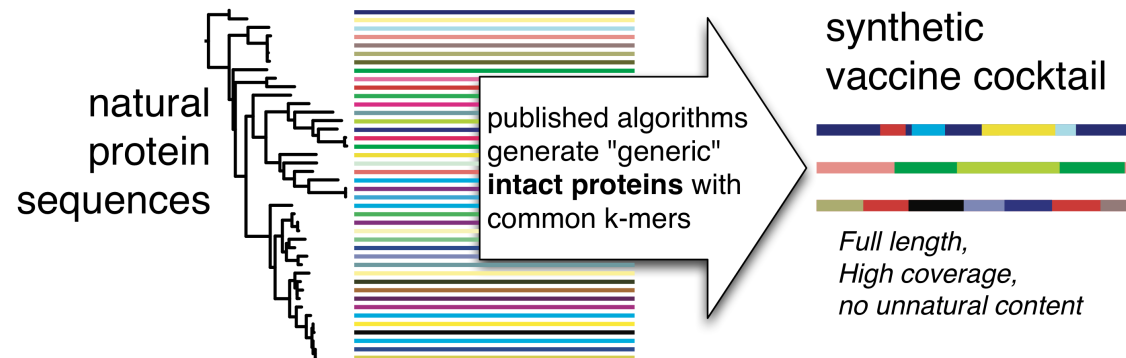
Allele	#	Start	End	Peptide Length	Sequence	Method used	Percentile Rank
HLA-B*15:01	1	6	13	8	DMNLPGRW	NetMHCpan	26
HLA-B*15:01	1	3	13	11	VLEDMNLPGRW	NetMHCpan	27
HLA-B*15:01	1	3	11	9	VLEDMNLPG	Consensus (ANN,SMM,CombLib_Sidney2008)	27.60
HLA-B*15:01	1	8	17	10	NLPGRWKPKM	NetMHCpan	31
HLA-B*15:01	1	7	17	11	MNLPGRWKPKM	NetMHCpan	35
HLA-B*15:01	1	2	9	8	TVLEDMNL	NetMHCpan	36
HLA-B*15:01	1	2	11	10	TVLEDMNLPG	NetMHCpan	47
HLA-B*15:01	1	4	11	8	LEDMNLPG	NetMHCpan	48

HIV epitopes are densely packed at the population level

- Vaccinating a diverse population with individual epitopes is infeasible
- Escape forms for one HLA are frequently sensitive for a different HLA
- It may not be necessary to *predict* epitopes — but only to *deliver* them
- Optimized immunogen cocktails could deliver most epitopes likely to be present in infecting virus



Vaccine Design Tools (Mosaic/Epigraph)



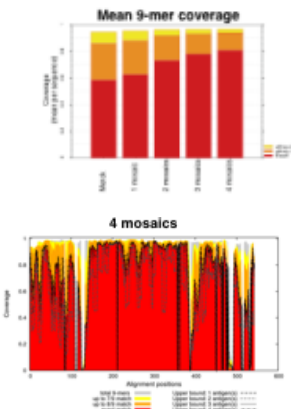
Design Tools

Generate candidate vaccine protein cocktails that optimize coverage of potential T-cell epitopes (as linear k -mers) based on frequencies in sets of natural pathogen sequences — “all-natural” throughout, including breakpoints

Mosaic Vaccine Designer — genetic algorithm (Fischer et al. 2007)

Epigraph — graph theoretic approach (Theiler et al. 2016)

Evaluation tools



Epitope Coverage Assessment (EPICOVER)

Alignment-independent “k-mer” coverage by vaccines or peptides.

Positional Epitope Coverage Assessment (POSICOVER)

Alignment-based coverage by vaccines or peptides.

Mosaic Vaccine Designer

Inputs

Target set: natural protein sequences from a diverse pathogen population (alignment optional).

Cocktail size: how many mosaic protein sequences to generate.

Epitope length: default is 9 amino-acids.

Method: genetic algorithm

Linear optimization: helpful for both T-cell and linear aspects of B-cell epitopes

Epitope length is transferable...

DEMONSTRATED EFFECTIVENESS

Improved immunogenicity

HIV, SIV, HCV, *Chlamydia*

Protection from challenge (non-human models):

SHIV, Influenza, FMDV, Ebola

Many human HIV trials in process

HIV sequence database

DATABASES SEARCH ALIGNMENTS TOOLS PUBLICATIONS GUIDES Search Site

Mosaic Vaccine Designer

Purpose: The Mosaic Vaccine Designer will generate candidate vaccine protein cocktails that optimize the coverage, by a small set of mosaic proteins that could be included in a vaccine cocktail, of potential T-cell epitopes in a large diverse set of proteins. The resulting 'mosaic' proteins in the proposed vaccine cocktail resemble real proteins from the input set of natural viral proteins (the 'training set'), but are assembled from fragments of the natural proteins using a genetic algorithm (a computational optimization method). This method was first applied to HIV, but is readily generalized and could be applied to other variable pathogens.

Functions:

- 'Create mosaic sequence cocktail' runs the genetic algorithm to generate a cocktail of synthetic sequences with near-optimal coverage
- 'Pick the best natural sequences' selects unmodified natural sequences from the training set in order of coverage
- 'See the coverage distribution of natural sequences' shows the coverages of a randomly selected set of natural sequence cocktails

Usage: Paste your protein sequences in the box below, or upload a file containing sequences. Most common [sequence formats](#) are accepted. As soon as your job is completed, a link to your results will be sent to your email address which you provided. To manage more detailed parameters, go to the Advanced Input. (Your job may take several hours or even days, according to your input.)

Related Programs:

- [Epitope Coverage Assessment Tool-Epicover](#)
- [Positional Epitope Coverage Assessment Tool-Posicover](#)

Reference: [Polyvalent vaccine design article](#) | [Pubmed version](#)

Input

Paste set of protein sequences
☒ Sample Input

```
A1.CM. .a
MGGNWSKSSLVGWPEIRERMRRAPPTPTTPAAKGVGAVSQDLAKHGAIT
A1.KE.99a
MGGKWSKSSIVGWPEVRRRIQOTPPAARGVGAVSQDLEKHGAITSSNINHS
A1.KE.99b
MGGIWSKRSTRGWSEVRERIRQTPPTPAARGVGAVSQDLARHGAVTSSNVN
```

Or upload protein sequence file

Options

Basic Advanced

Function

☒ Create mosaic sequence cocktail
☐ Pick the best natural sequences
☐ See the coverage distribution of natural sequences

Cocktail Size (1-10)

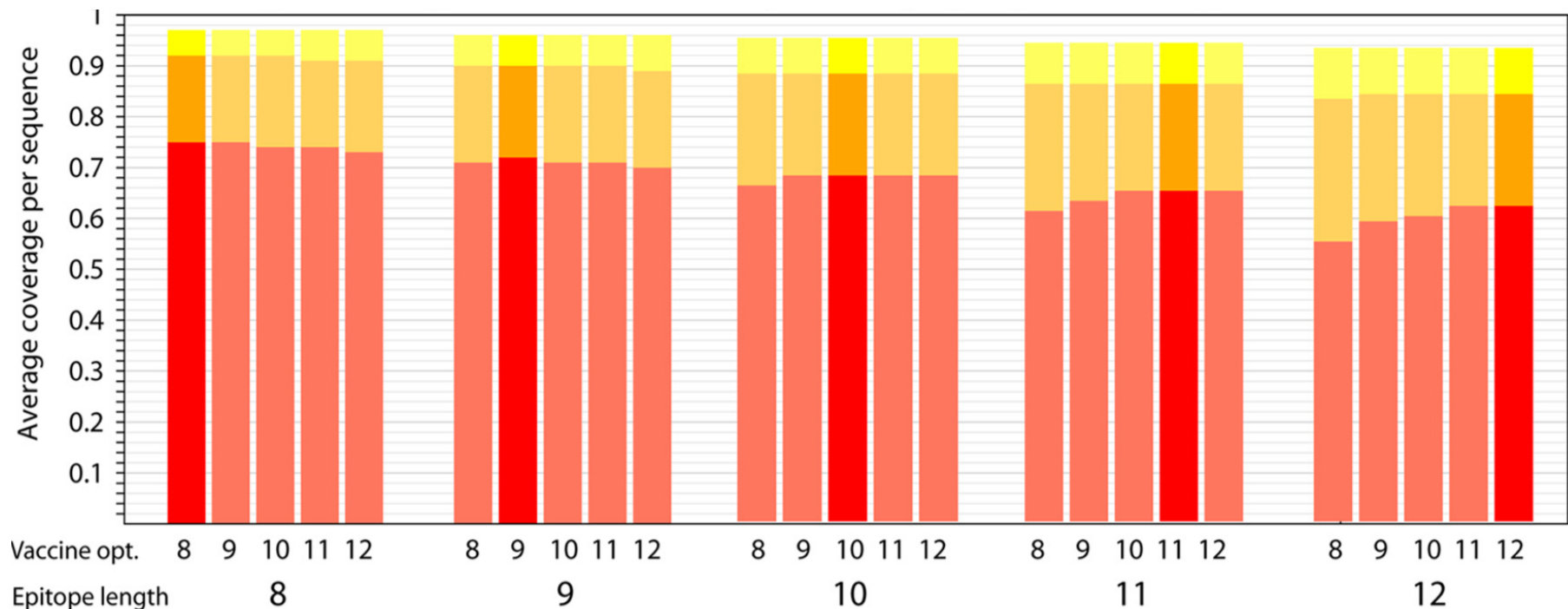
Epitope Length (8-12)

Rare Threshold

Paste fixed sequences

Or upload fixed sequence file

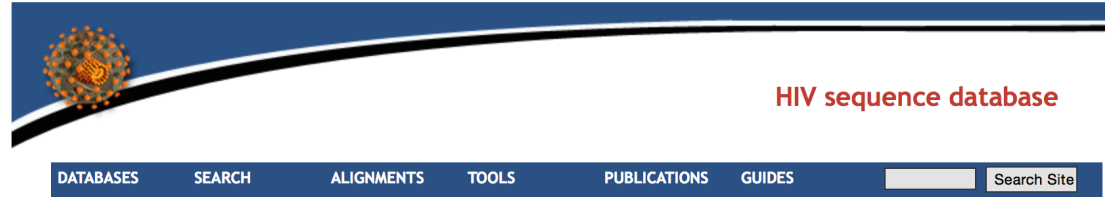
***k*-mer coverage is relatively stable for different values of *k* (potential epitope lengths)**



In other words, optimizing for potential CD8+ T-cell epitopes ($k=9$) yields good coverage of potential CD4+ T-cell epitopes ($k=12$), too.

[Korber et al., 2009] T-cell vaccine strategies for human immunodeficiency virus, the virus with a thousand faces. J Virol, 83(17):8300–14.

EPIGRAPH



Inputs

Target set: natural protein sequences for the pathogen population (alignment optional).

Cocktail size: how many mosaic proteins in the output set.

Epitope length: default is 9 amino-acids.

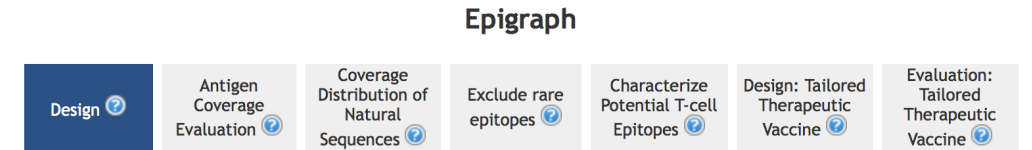
Method: evaluation of acyclic graph

Advantages over mosaic

Essentially optimal (fractionally better coverage)

Much faster: allows iteration and comparison of multiple input sets and alternate designs

Reference: Theiler, J., Yoon, H., Yusim, K., Picker, L. J., Fruh, K., and Korber, B. (2016). Epigraph: A vaccine design tool applied to an HIV therapeutic vaccine and a pan-filovirus vaccine. *Sci Rep*, 6:33987.



Input

Protein sequences
[Sample Input]

Choose File no file selected

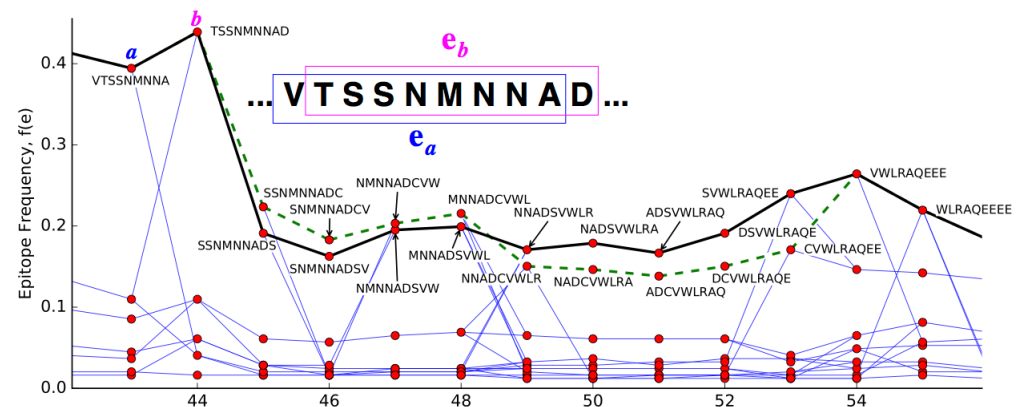
Options

Algorithm ☒ Unaligned sequence algorithm (Gaps will be removed before calculation if seqs are aligned) ☐ Aligned sequence algorithm (Seqs must be the same length. If not, automatically "padded")

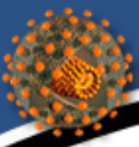
Epitope length

of seqs in vaccine pool

Email results ☐



EPIGRAPH — exclude rarities



HIV sequence database

DATABASES SEARCH ALIGNMENTS TOOLS PUBLICATIONS GUIDES Search Site

Epigraph

Design ? Antigen Coverage Evaluation ? Coverage Distribution of Natural Sequences ? **Exclude rare epitopes ?** Characterize Potential T-cell Epitopes ? Design: Tailored Therapeutic Vaccine ? Evaluation: Tailored Therapeutic Vaccine ?

Input

Protein sequences

no file selected

or

Epigraph design job number

Options

Epitope length

of sequences in vaccine

Email results ☐

Epitope Coverage Assessment - Epicover

Inputs:

1. Vaccine set

2. Test set (target sequences)

Can report on subsets defined according to the first several characters in sequence names or user-defined subsets

Input

Use output from MakeVaccine tool

Provide a job number to access output from the [Mosaic Vaccine Designer](#) tool:

OR

Provide input sequences

Paste antigen protein sequence(s):
[\[Sample Input\]](#)

upload more [+] antigen sequence files

and/or upload as files:

Paste test set protein sequences:

upload more [+] test sequence files

and/or upload as files:

Options

Send results as an email instead of displaying in browser
(useful in case of a browser time-out): ☐

Nominal epitope length:

Maximum amino acid mismatches to score (range from 0):

Minimum number of occurrences of a potential epitope
in viral protein set to consider for coverage:

Precision to use when reporting coverage: decimal places

Advanced Options

Upload file of grouped sequence names

Report on subsets defined according to first character(s) in sequence names

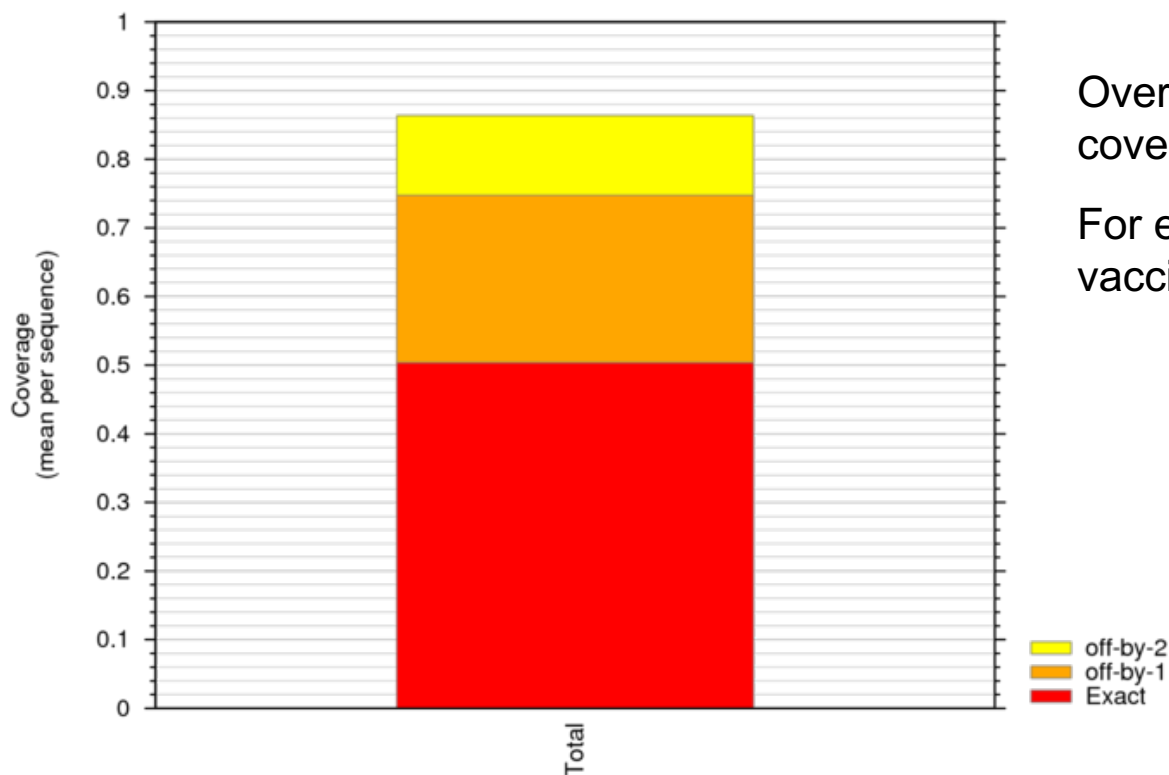
Epicover output (mean coverage per sequence)

vaccine set	subset	subset count	Off-by-0	Off-by-1	Off-by-2	rare(<3,>1)	unique	absent
vaccine_set_from_user	Total	39	0.5037	0.7474	0.8636	91	61	38
vaccine_set_from_user	A	6	0.4294	0.7086	0.8417	7	1	38
vaccine_set_from_user	B	4	0.7263	0.8911	0.9460	44	23	38
vaccine_set_from_user	C	4	0.5786	0.8449	0.9602	47	37	38
vaccine_set_from_user	D	4	0.5764	0.8268	0.9218	12	0	38
vaccine_set_from_user	F	8	0.4821	0.7316	0.8786	2	0	38
vaccine_set_from_user	G	4	0.4578	0.7126	0.8367	5	0	38
...

Overall summaries of *k*-mer coverage

Epcover output (mean coverage per sequence)

vaccine set	subset	subset count	Off-by-0	Off-by-1	Off-by-2	rare(<3,>1)	unique	absent
vaccine_set_from_user	Total	39	0.5037	0.7474	0.8636	91	61	38
vaccine_set_from_user	A	6	0.4294	0.7086	0.8417	7	1	38
vaccine_set_from_user	B	4	0.7263	0.8911	0.9460	44	23	38
vaccine_set_from_user	C	4	0.5786	0.8449	0.9602	47	37	38
vaccine_set_from_user	D	4	0.5764	0.8268	0.9218	12	0	38
vaccine_set_from_user	F	8	0.4821	0.7316	0.8786	2	0	38
vaccine_set_from_user	G	4	0.4578	0.7126	0.8367	5	0	38
...

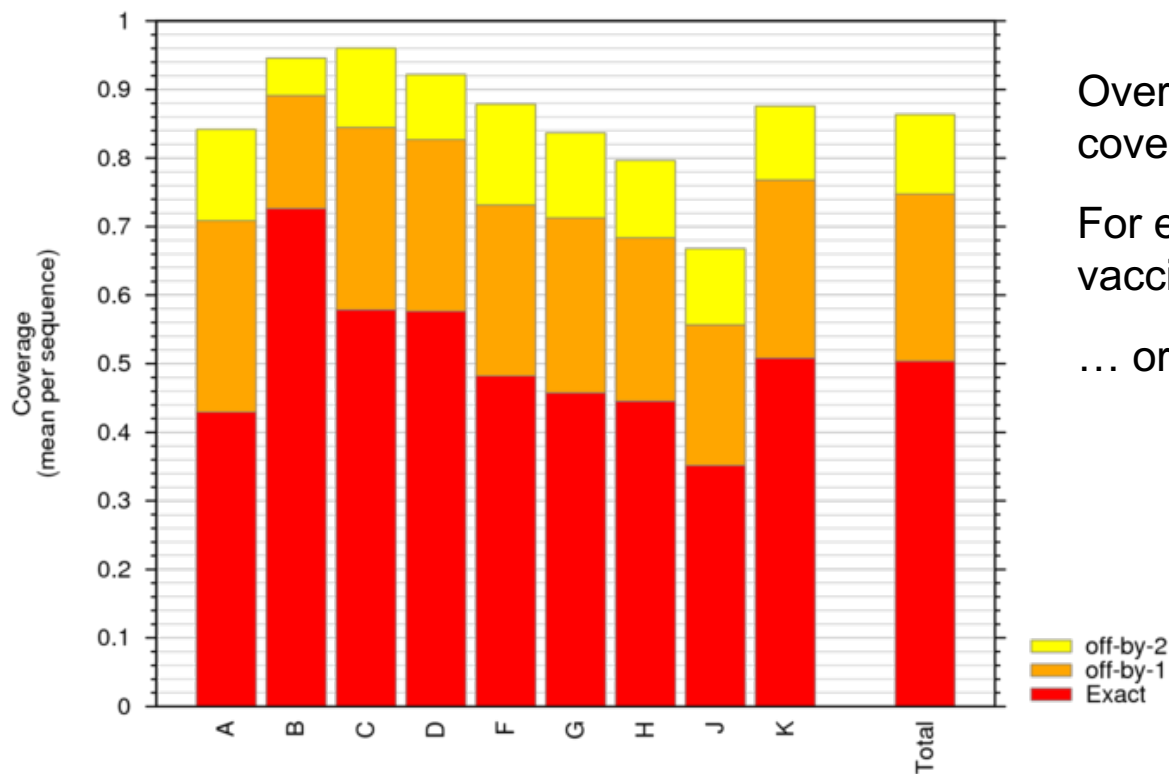


Overall summaries of *k*-mer coverage

For entire set (to compare with other vaccine candidates)

Epicover output (mean coverage per sequence)

vaccine set	subset	subset count	Off-by-0	Off-by-1	Off-by-2	rare(<3,>1)	unique	absent
vaccine_set_from_user	Total	39	0.5037	0.7474	0.8636	91	61	38
vaccine_set_from_user	A	6	0.4294	0.7086	0.8417	7	1	38
vaccine_set_from_user	B	4	0.7263	0.8911	0.9460	44	23	38
vaccine_set_from_user	C	4	0.5786	0.8449	0.9602	47	37	38
vaccine_set_from_user	D	4	0.5764	0.8268	0.9218	12	0	38
vaccine_set_from_user	F	8	0.4821	0.7316	0.8786	2	0	38
vaccine_set_from_user	G	4	0.4578	0.7126	0.8367	5	0	38
...



Overall summaries of *k*-mer coverage

For entire set (to compare with other vaccine candidates)

... or by pathogen subset

Positional Epitope Coverage Assessment - Posicover

Input

Provide a job # from [Mosaic Vaccine Designer](#): (Only the antigen set is used. Provide the ALIGNED viral

AND/OR

Paste antigen protein set or peptide cocktail: (alignment not required) [\[Sample Input\]](#)

upload more [+] antigen files

and/or upload antigen file(s): No file selected.

Test set proteins

Paste **ALIGNED** test viral protein set: [\[Sample Input\]](#)

or upload an **ALIGNED** test proteins file: No file selected.

Options

Nominal epitope length:

Antigen counts to compute upper bounds:

Plots to make

Hits in their natural positions ☒

Misses in their natural positions ☒

Hits and misses in their natural positions ☒

Hits ranked by coverage ☒

Misses ranked by coverage ☒

N-mer coverage by positions ☒

Ranked n-mer coverage ☒

Alignment Thumbnail ☒

N-mer coverage directly on alignment ☒

- INPUTS

1. Vaccine/peptide sequences
2. ALIGNED target set

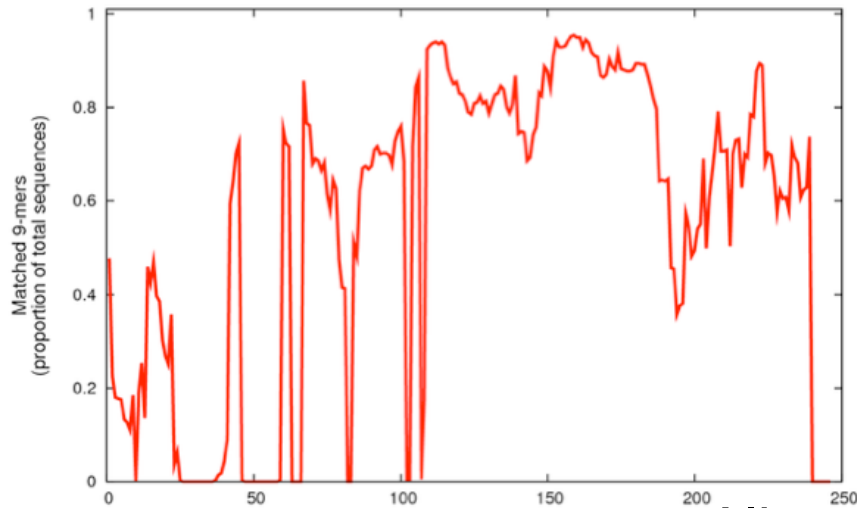
- OUTPUTS

1-dimensional (by alignment column)

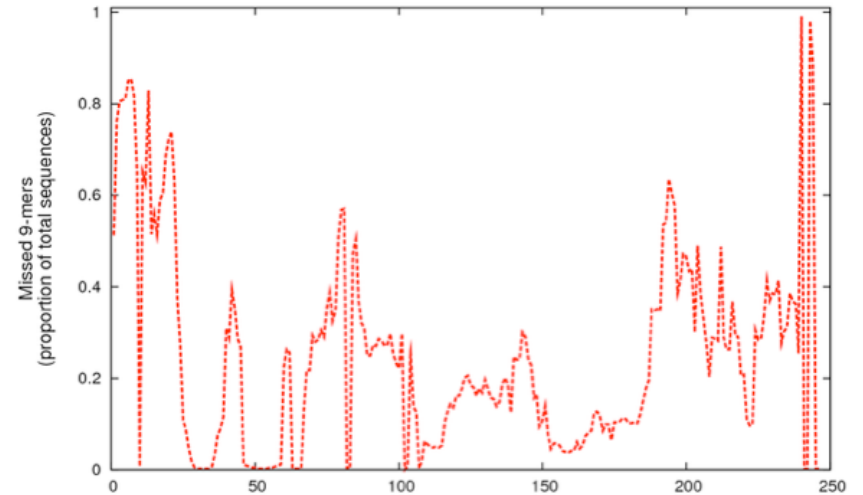
2-dimensional (by sequence and alignment column)

Posicover output (1-dimensional summaries)

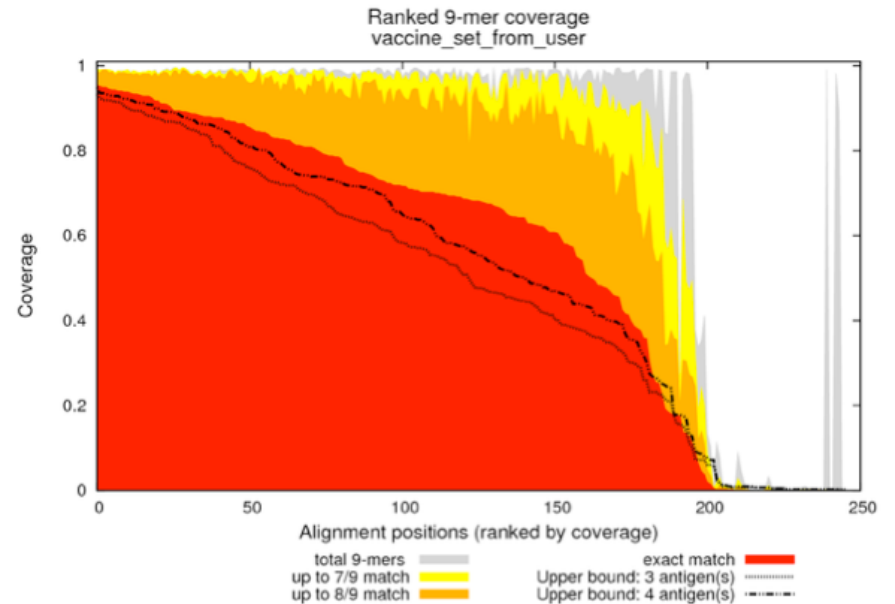
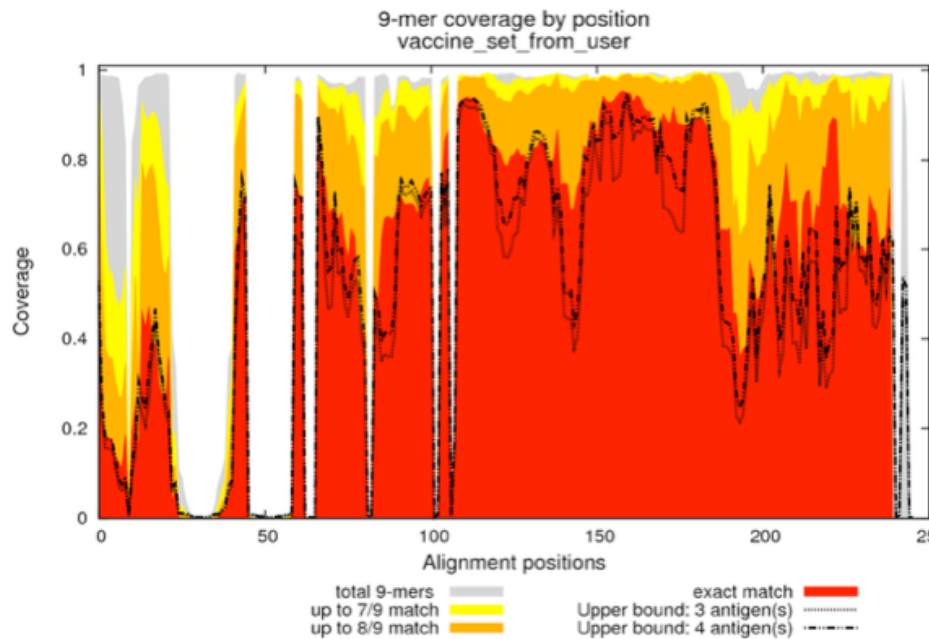
Matched 9-mers



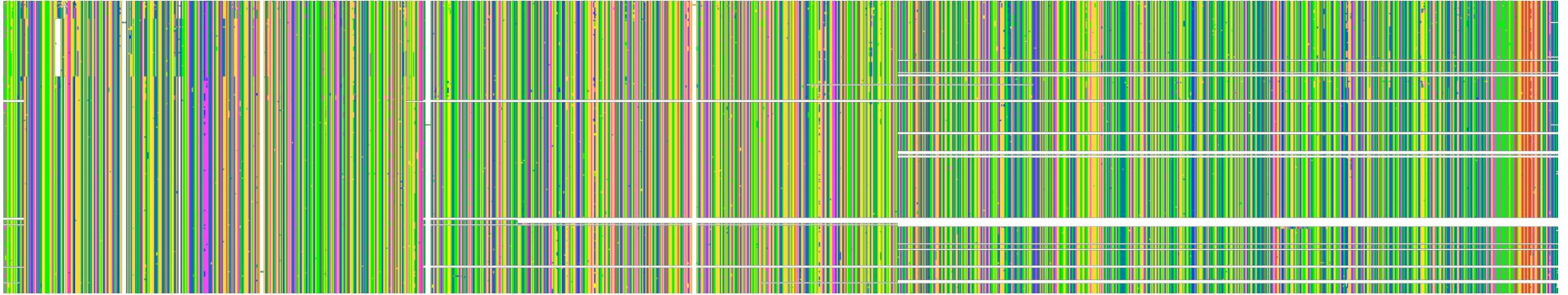
Missed 9-mers



Alignment positions



Posicover output (2 dimensional)

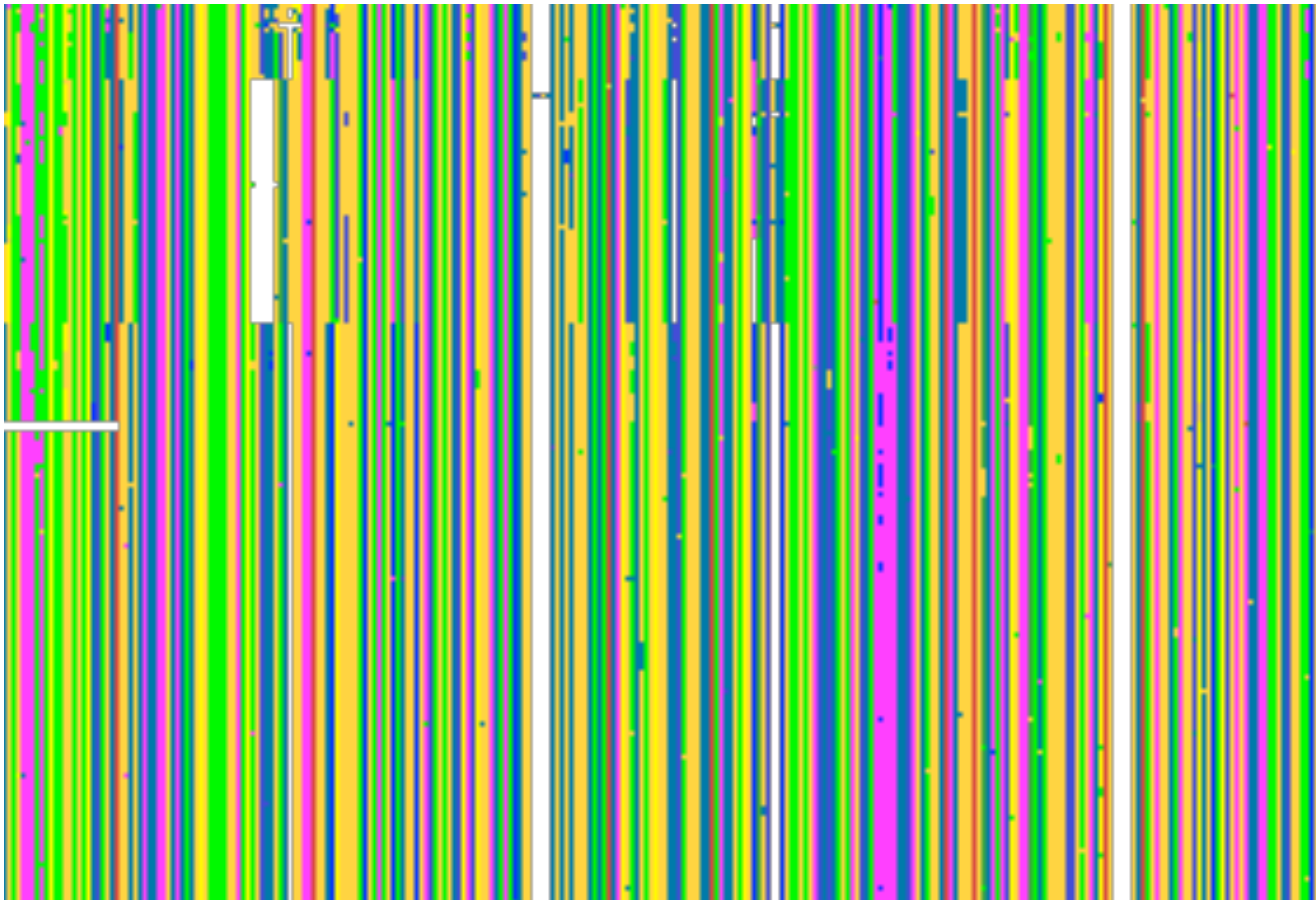


Pixel-based Alignment view

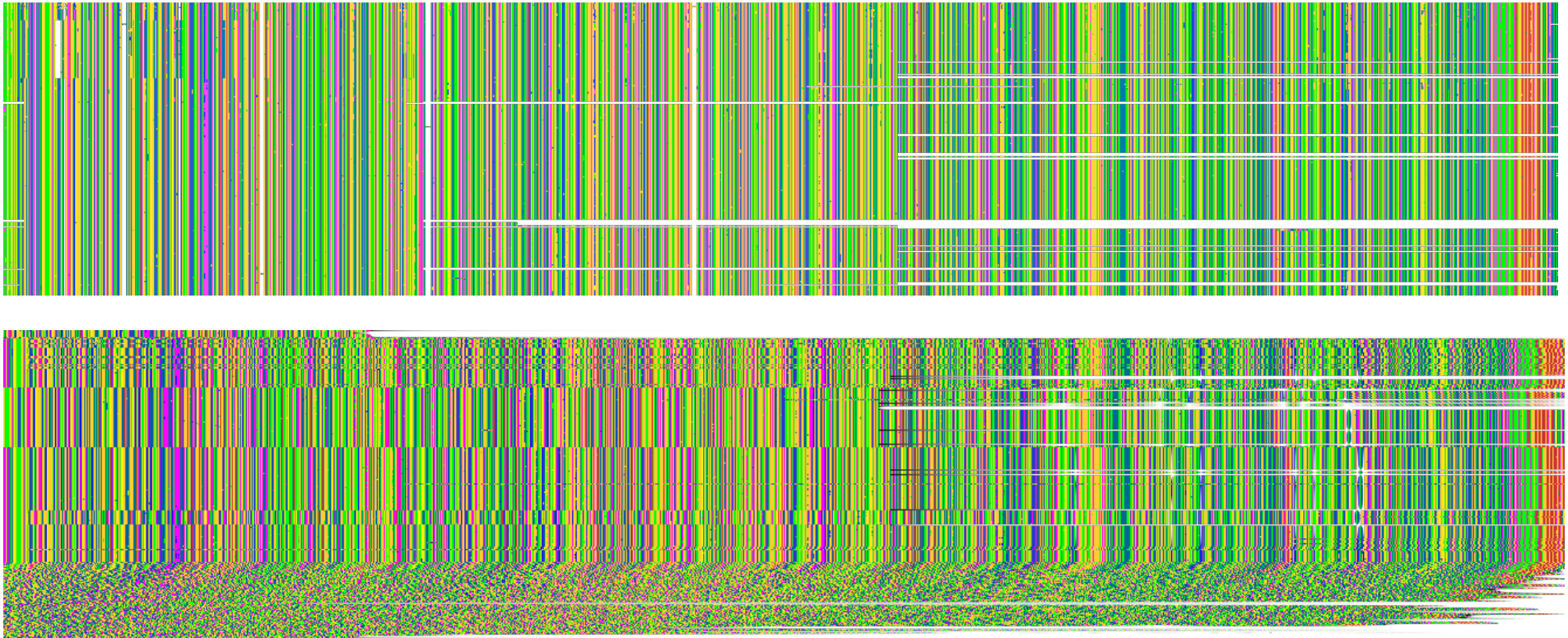
Each amino-acid represented as a single-colored square

Allows quick detection of gross errors in alignment

Posicover output (2 dimensional)



Posicover output (2 dimensional)

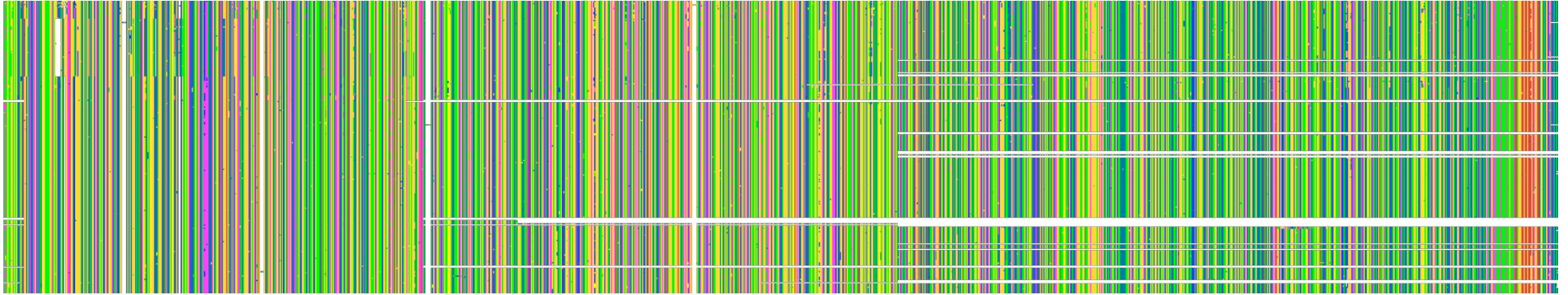


Pixel-based Alignment view

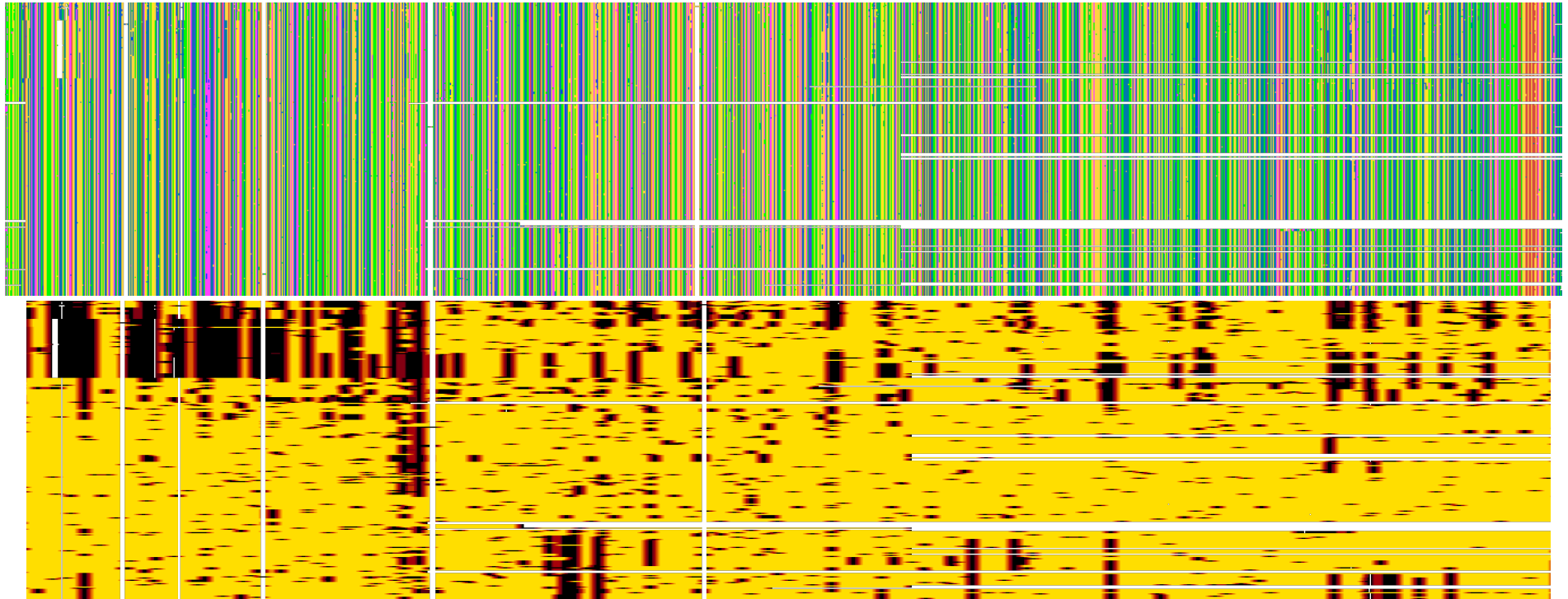
Each amino-acid represented as a single-colored square

Allows quick detection of gross errors in alignment

Posicover output (2 dimensional)

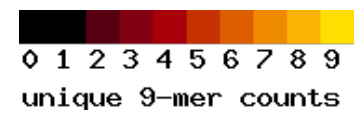
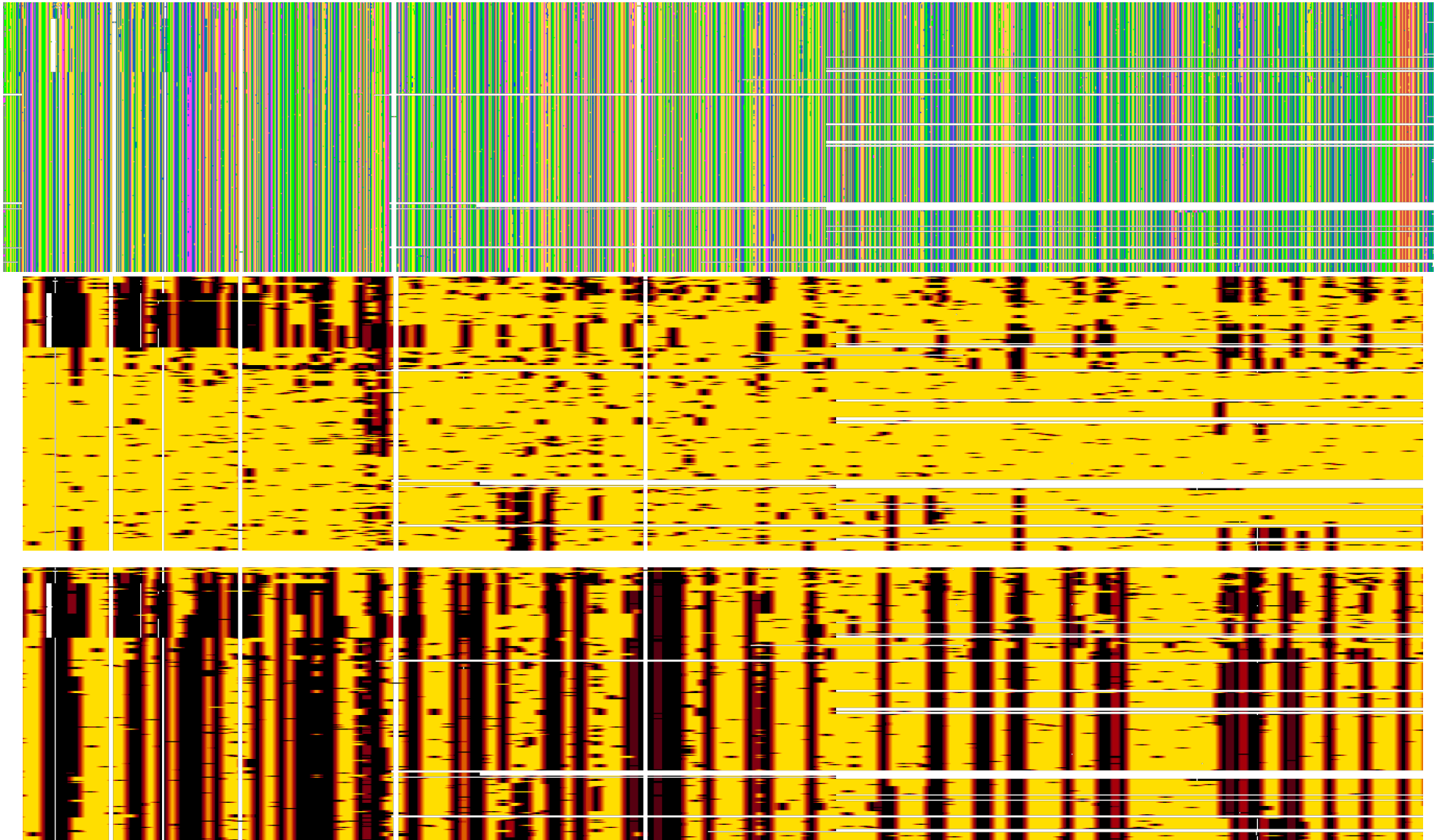


Posicover output (2 dimensional)

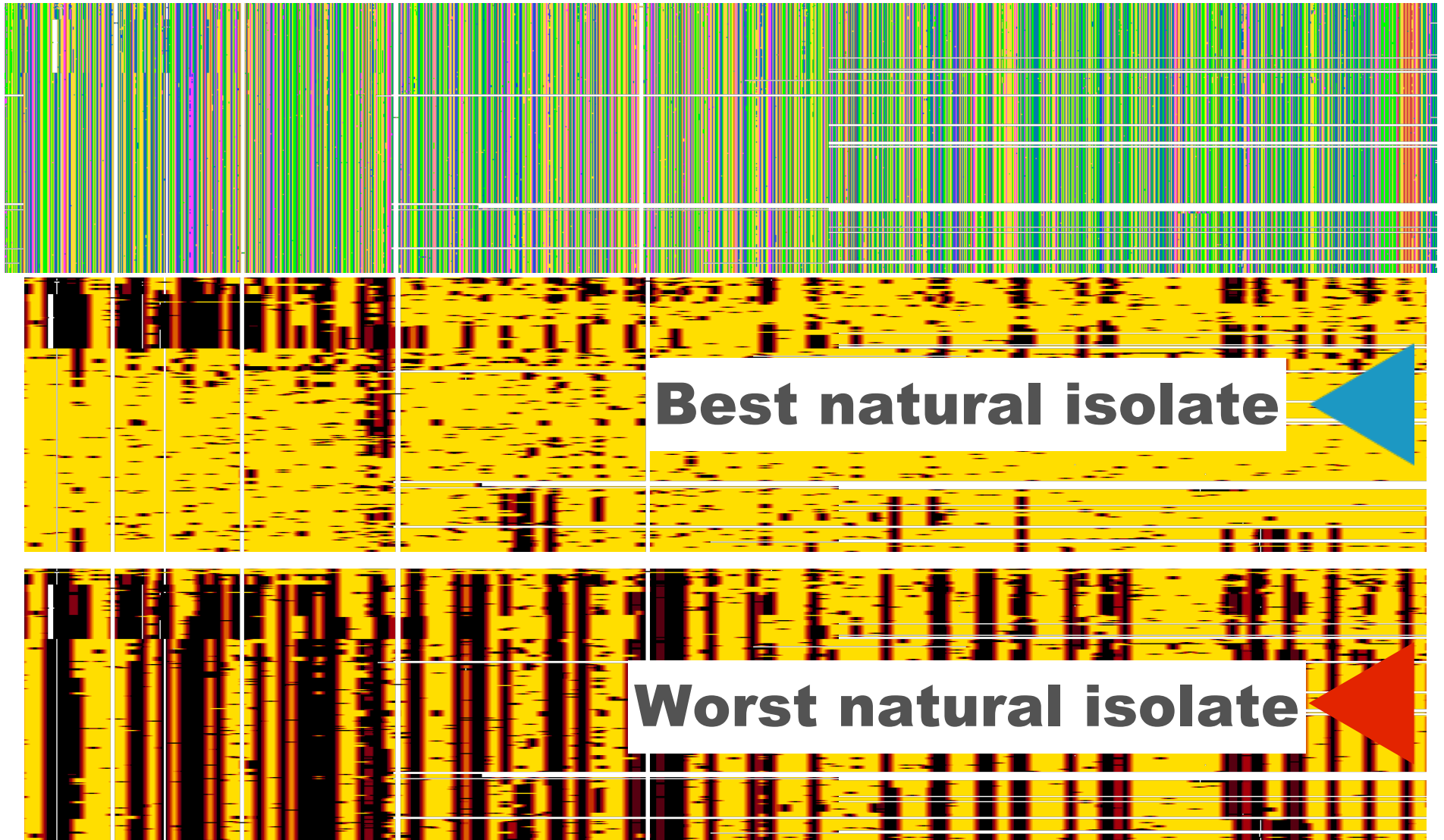


POSICOVER *K*-MER COVERAGE
(YELLOW-BLACK GRADIENT SHOWS HOW MANY OF EACH
RESIDUE'S *K*-MERS APPEAR IN VACCINE)

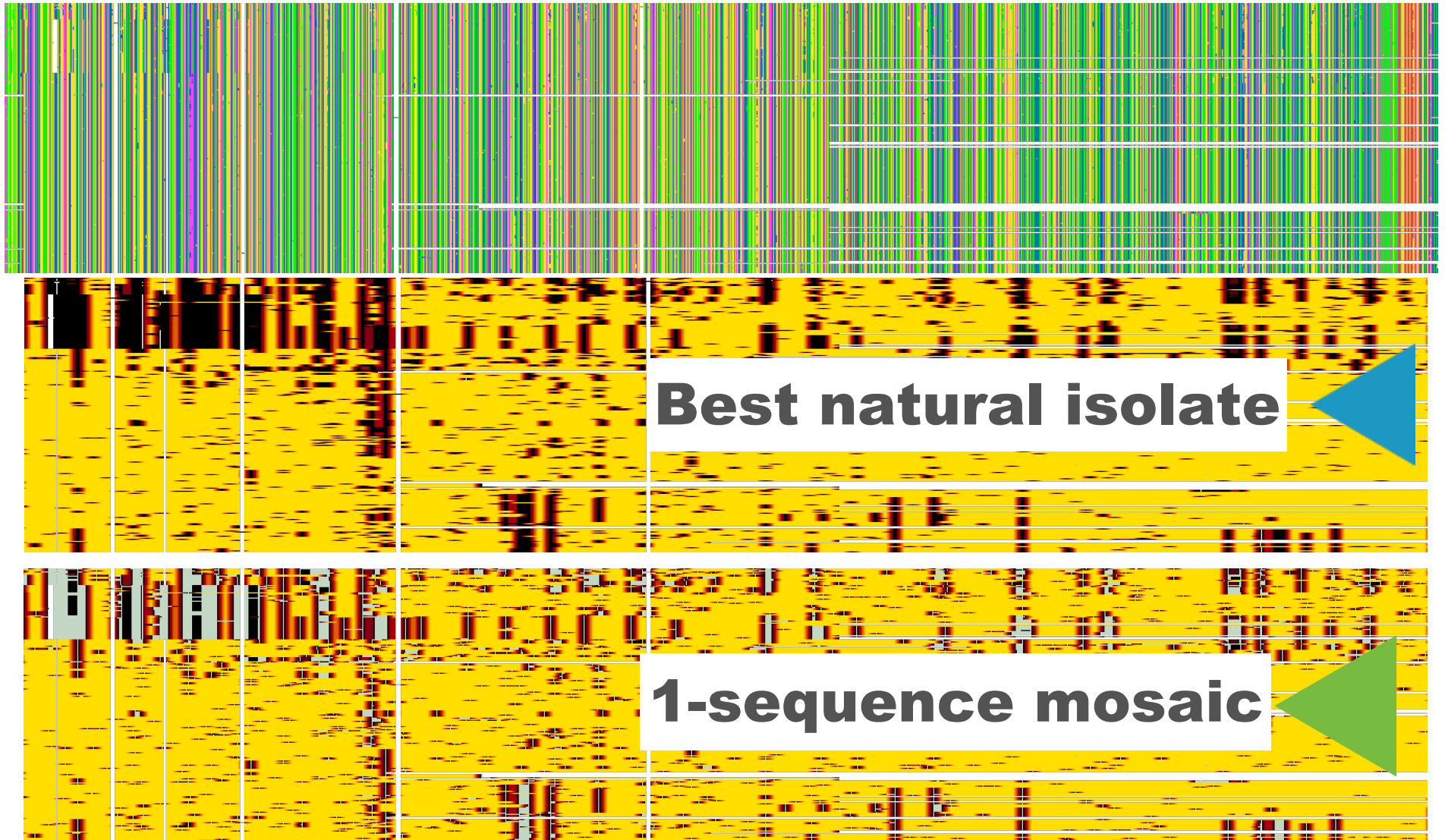
Posicover output (2 dimensional)



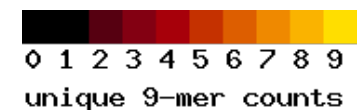
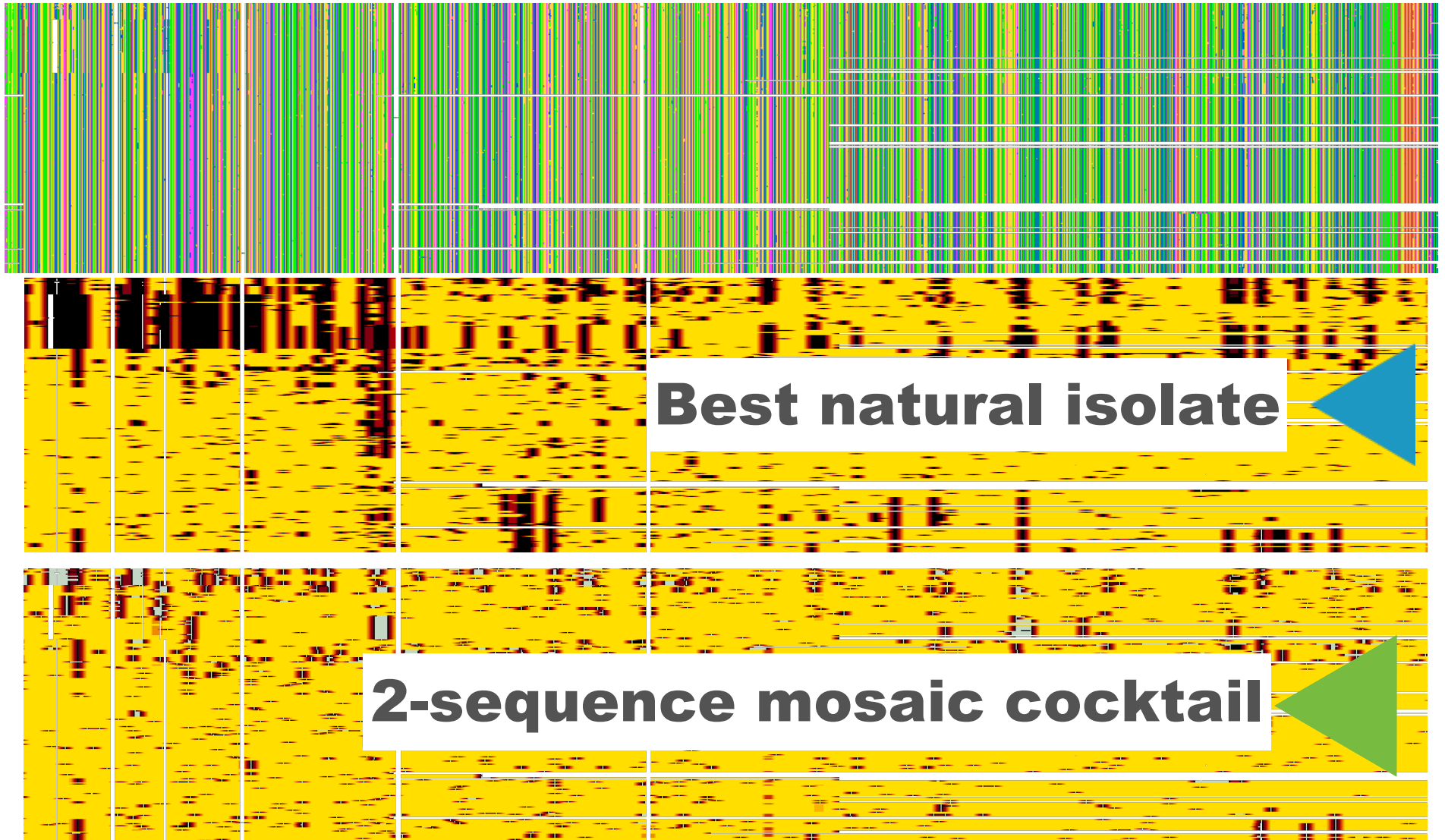
Posicover output (2 dimensional)



Posicover output (2 dimensional)



Posicover output (2 dimensional)



Thank you for attending!

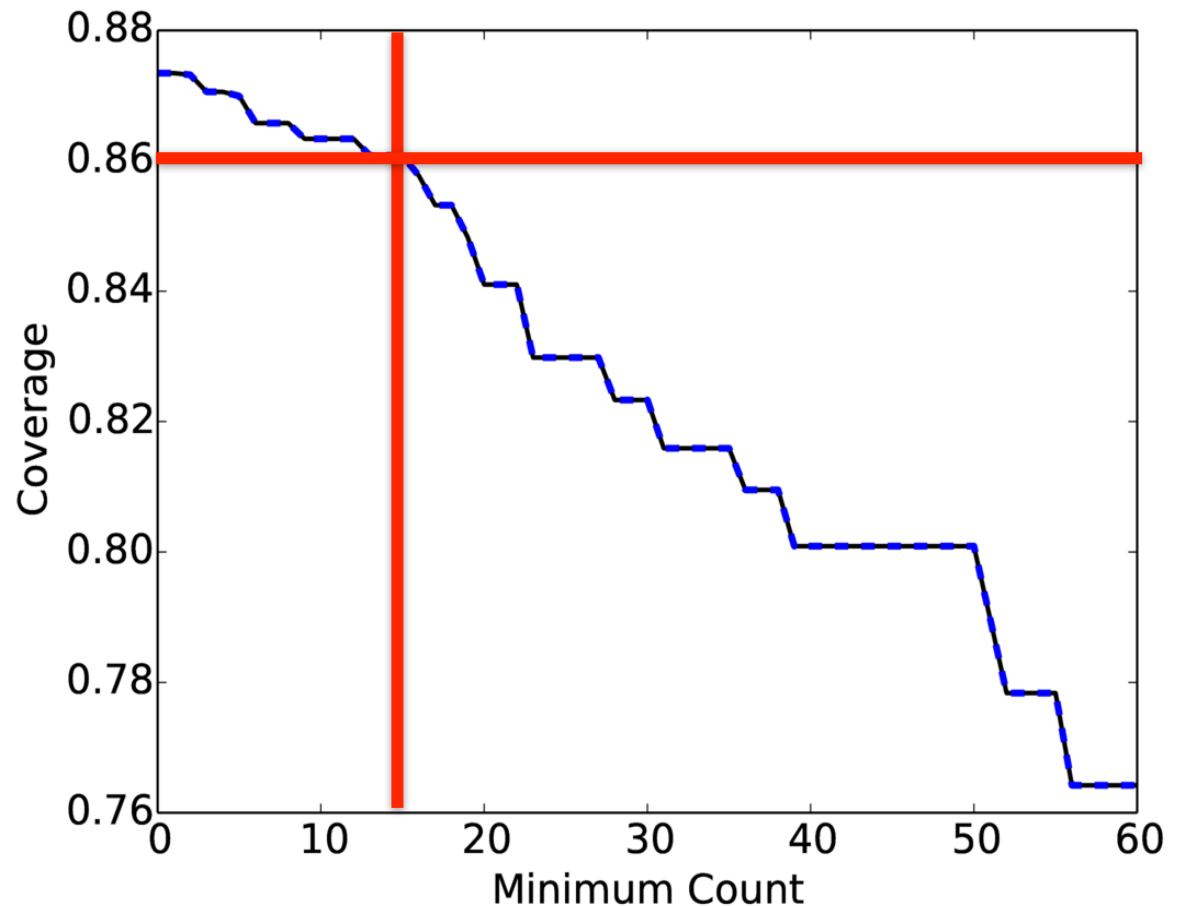
Please fill out our evaluation form!

Your comments will help us provide future training.

Contact us: seq-info@lanl.gov or immuno@lanl.gov

EPIGRAPH — exclude rarities

- Including *only* k-mers above an occurrence threshold drops coverage, but reducing responses to rare epitopes may be helpful.



Here, including only 9-mers that occur at least 14 times drops coverage very little.